

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:38:15 ; Search time 99 Seconds
(without alignments)
117.015 Million cell updates/sec

Title: US-09-856-070b-29

Perfect score: 193

Sequence: 1 EREKQMPREKEELMLRLD.....BEKTKARELSEQIQRALQ 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	191	99.0	436	4	AAG73954	Aag73954 Human col
2	191	99.0	579	7	ADC31630	Adc31630 Human nov
3	191	99.0	585	7	AD663987	Ad663987 Human pro
4	191	99.0	586	2	AAV27443	Aav27443 Amino aci
5	191	99.0	611	7	ADC31629	Adc31629 Human nov
6	191	99.0	622	4	AAU30004	Aau30004 Novel hum
7	191	99.0	628	7	ADC31631	Adc31631 Human nov
8	191	99.0	635	3	AA53356	Aab53356 Human col
9	156	80.8	34	4	AA882020	Aab82020 Human hep
10	138	71.5	52	4	AAU33060	Aau33060 Novel hum
11	129	66.8	576	7	AD254611	Ad254611 Human pro
12	129	66.8	576	7	AD45108	Ad45108 Human pro
13	129	66.8	577	6	AB859727	Ab859727 Human moe
14	128	66.8	577	7	AD870362	Ad870362 Moesin SE
15	129	66.8	577	7	AD77532	Ad77532 Human moe
16	121	62.7	593	4	ABG19947	Abg19947 Novel hum
17	120	62.2	583	7	AD879911	Ad879911 Mouse put
18	120	62.2	583	7	ADC56738	Adc56738 Murine ra
19	119	61.7	503	4	ABG16577	Abg16577 Novel hum
20	119	61.7	583	6	ABU89709	Abu89709 Protein d
21	117	60.6	579	7	ADC77529	Adc77529 Mouse moe
22	117	60.6	579	7	ADC77526	Adc77526 Zebrafish
23	79	40.9	107	4	ABU53290	Abu53290 Human tes
24	79	40.9	109	4	ABU53269	Abu53269 Human tes
25	79	40.9	109	4	ABU53270	Abu53270 Human tes

26	79	40.9	110	4	ABU53286	Abu53286 Human tes
27	79	40.9	115	4	ABU53289	Abu53289 Human tes
28	79	40.9	116	4	ABU53283	Abu53283 Human tes
29	79	40.9	282	4	ABU53282	Abu53282 Human tes
30	79	40.9	357	4	ABU53268	Abu53268 Human tes
31	79	40.9	2707	2	AAW27161	.AAW27161 Mouse rec
32	77	39.9	97	4	ABU53276	Abu53276 Human tes
33	77	39.9	120	4	ABU53279	Abu53279 Human tes
34	77	39.9	135	4	ABU53281	Abu53281 Human tes
35	77	39.9	339	4	ABU53271	Abu53271 Human tes
36	75	38.9	105	4	ABU53274	Abu53274 Human tes
37	73.5	38.1	650	7	ADB65482	Adb65482 Human pro
38	73	37.8	730	4	ABB71879	Abb71879 Drosophil
39	73	37.8	1898	2	AA30795	Aay30795 A human t
40	73	37.8	1898	7	ADD48869	Add48869 Human pro
41	72.5	37.6	85	4	ABU53273	Abu53273 Human tes
42	72.5	37.6	105	4	ABU53277	Abu53277 Human tes
43	72.5	37.6	144	4	ABU53278	Abu53278 Human tes
44	71.5	37.0	2274	4	ABB58657	Abb58657 Drosophil
45	71	36.8	152	4	ABU53272	Abu53272 Human tes

ALIGNMENTS

RESULT 1
AAG73954
ID AAG73954 standard; protein; 436 AA.
XX
AC AAG73954;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:4718.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US026524.
XX
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
(HUMA-) HUMAN GENOME SCI INC.
XX
Ruben SM, Barash SC, Birse CE, Rosen CA;
WPI; 2001-235357/24.
XX
N-PSDB; AAH33385.
XX
Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
useful for preventing, diagnosing and/or treating colorectal cancers.
XX
Claim 11; Page 6520-6521; 9803pp; English.
XX
AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
cancer-associated nucleic acid molecules (N) and proteins (P), where the
proteins are collectively known as colon cancer antigens. The colon
cancer antigens have cytostatic activity and can be used in gene therapy
and vaccine production. N and P may be used in the prevention, diagnosis
and treatment of diseases associated with inappropriate P expression. For
example, N and P may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of P by expressing inactive proteins or to
supplement the patients own production of P. Additionally, N may be used
to produce the colon cancer-associated ps, by inserting the nucleic acids
into a host cell and culturing the cell to express the proteins. N and P

CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027
 CC to 1052, 7921 and 7922
 CC
 XX Sequence 436 AA;
 SQ

Query Match 99.0%; Score 191; DB 4; Length 436;
 Best Local Similarity 97.6%; Pred. No. 2.7e-13;
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EREKQWREKEELMLRLQDXEETKKAERELSEQIORALQ 41
 DB 184 EREKQWREKEELMLRLQDYEEKTKKAERELSEQIORALQ 224

RESULT 2
 ADC31630
 ID ADC31630 standard; protein; 579 AA.
 AC ADC31630;
 XX
 XX 18-DEC-2003 (first entry)
 DT Human novel polypeptide sequence, SEQ ID NO:1712.
 DE
 DE Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; anti-anaemic; anticoagulant; thrombolytic; vulnary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 6q25.2-26.
 XX
 XX Homo sapiens.
 OS
 XX WO2003029271-A2.
 FN
 XX 10-APR-2003.
 PD
 XX 24-SEP-2002; 2002WO-US030474.
 PF
 XX 24-SEP-2001; 2001US-0324631P.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 FI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 FI Haley-Vicente D, Drmanac RT;
 XX
 XX WPI; 2003-371981/35.
 DR
 XX N-ESDB; ADC30659.
 DR
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 FT treating conditions such as neurodegenerative diseases, anemias, platelet
 FT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 FT cancer.
 FT
 XX Claim 20; SEQ ID NO 1712; 1185pp; English.
 PS
 XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or

CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancers. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 579 AA;
 SQ

Query Match 99.0%; Score 191; DB 7; Length 579;
 Best Local Similarity 97.6%; Pred. No. 3.7e-13;
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EREKQWREKEELMLRLQDXEETKKAERELSEQIORALQ 41
 DB 302 EREKQWREKEELMLRLQDYEEKTKKAERELSEQIORALQ 342

RESULT 3
 ADE63987
 ID ADE63987 standard; protein; 585 AA.
 XX
 XX ADE63987;
 AC
 XX 29-JAN-2004 (first entry)
 DT Human Protein P15311, SEQ ID NO 9933.
 DE
 XX Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 KW
 XX Homo sapiens.
 OS
 XX WO2003016475-A2.
 FN
 XX 27-FEB-2003.
 PD
 XX 14-AUG-2002; 2002WO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR
 XX 01-NOV-2001; 2001US-0346382P.
 PR
 XX 26-NOV-2001; 2001US-0333347P.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PA
 PA Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 DR GENBANK; P15311.
 DR
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PT
 PS Claim 1; Page; 1017pp; English.
 PS
 XX

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 585 AA;

Query Match 99.0%; Score 191; DB 7; Length 585;
Best Local Similarity 97.8%; Pred. NO. 3.7e-13;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EREKEQMMREKEELMLRLQDXEETKKAERELSEQIORALQ 41
|||||
Db 333 EREKEQMMREKEELMLRLQDYEEKTKKAERELSEQIORALQ 373

RESULT 4
AA27443
ID AAY27443 standard; protein; 586 AA.

AC AAY27443;

XX 26-NOV-1999 (first entry)

XX Amino acid sequence of human ezrin polypeptide.

XX Pharmaceutical; ezrin; mutant; tumor; metastasis; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT FT Misc-difference 354

FT /note= "the Tyr at this position can be mutated

FT (preferably to a Phe) to construct an ezrin mutant of the

XX invention"

XX WO9947150-A2.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-BP002054.

XX 18-MAR-1998; 98US-00040725.

XX (CURI-) INST CURIE.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Arpin M, Crepaldi T, Gautreau A, Louvard D;

XX WPI; 1999-561851/47.

XX New composition for prevention and treatment of tumors and metastasis.
PT Example 1; Fig 1; 31pp; English.

XX The invention provides a pharmaceutical composition containing ezrin
CC protein, RNA or DNA mutated on tyrosine 353, or a functional fragment or
CC derivative of the ezrin mutant. The new composition is useful for
CC prevention and/or treatment of tumors, and especially metastasis. The
CC present sequence represents the amino acid sequence of human ezrin
CC (before the maturation by deletion of the first amino acid Met)
XX SQ Sequence 586 AA;

Query Match 99.0%; Score 191; DB 2; Length 586;
Best Local Similarity 97.6%; Pred. No. 3.7e-13;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EREKEQMMREKEELMLRLQDXEETKKAERELSEQIORALQ 41
|||||
Db 334 EREKEQMMREKEELMLRLQDYEEKTKKAERELSEQIORALQ 374

RESULT 5

ADC31629

ID ADC31629 standard; protein; 611 AA.

XX AC ADC31629;

XX 18-DEC-2003 (first entry)

XX Human novel polypeptide sequence, SEQ ID NO:1711.

XX Human; diagnostic; drug screening; forensics; gene mapping;

XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;

XX neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

XX ulcers; osteoporosis; autoimmune disease; cancer;

XX molecular weight marker; food supplement; antiparkinsonian; nootropic;

XX neuroprotective; anti-naemic; anticoagulant; thrombolytic; vulnerary;

XX antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

XX gene therapy; chromosome 6q25.2-26.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

XX Haley-Vicente D, Drmanac RT;

XX WPI; 2003-371981/35.

XX DR N-PSDB; ADC30658.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or

XX treating conditions such as neurodegenerative diseases, anemias, platelet

XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

XX cancer.

XX Claim 20; SEQ ID NO 1711; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-

XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The

XX invention also relates to nucleic acid sequences over 99% identical with

XX the novel human cDNAs. The invention additionally encompasses expression

XX vectors and host cells comprising a nucleic acid of the invention; the

Claim 20; Page 219; 765pp; English.

The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell proliferation; to regulate haematopoiesis; and in bone, cartilage tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid sequences of novel human secreted proteins of the invention

```

an
SQ      Sequence 622 AA;

Query Match      99.0%; Score 191; DB 4; Length 622;
Best Local Similarity 97.6%; Pred. No. 4e-13;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 BREKQNMREKEELMLRLQDXEETKKAERELSEQIORALQ 41

370 BREKQNMREKEELMLRLQDYEEETKKAERELSEQIORALQ 410

RESULT 7
ADC31631
ID ADC31631 standard: protein: 628 AA.

DT	18-DEC-2003 (first entry)
XX	
DE	Human novel polypeptide se

XX	Human; diagnostic; drug screening; forensics; gene mapping;
XX	biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW	neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW	ulcers; osteoporosis; autoimmune disease; cancer;
KW	molecular weight marker; food supplement; antiparkinsonian; neurotropic;
KW	neuroprotective; antinaemic; anticoagulant; thrombolytic; vulnerary;
KW	antulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW	gene therapy; chromosome 6q25.2-26.
XX	
OS	Homo sapiens.
XX	
FN	WC003029271-A2.
XX	
PD	10-APR-2003.
XX	
PF	24-SEP-2002; 2002WO-US030474.
XX	
PR	24-SEP-2001; 2001US-0324631P.
XX	
PA	(HYSEQ-) HYSEQ INC.
XX	
PI	Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI	Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI	Haley-Vicente D, Drmanac RT;
XX	
XX	WPI; 2003-371981/35.
DR	DR N-PSDB: ADC30660.

PT treating conditions such as neurodegenerative diseases, anemias, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer.


```
XX PS Claim 20; SEQ ID NO 1713; 1185pp; English.
XX CC The invention relates to 971 novel human cDNA sequences (ADC29919-
XX CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX CC invention also relates to nucleic acid sequences over 99% identical with
XX CC the novel human cDNAs. The invention additionally encompasses expression
XX CC vectors and host cells comprising a nucleic acid of the invention; the
XX CC recombinant production of a polypeptide of the invention; an antibody
XX CC against a polypeptide of the invention; a method of detecting
XX CC polynucleotides or polypeptides of the invention; and methods of
XX CC identifying a compound which binds to a polypeptide of the invention. The
XX CC invention further discloses methods of preventing, treating or
XX CC ameliorating a medical condition; kits comprising polynucleotide probes
XX CC and/or monoclonal antibodies for carrying out the methods of the
XX CC invention; methods for the identification of compounds that modulate the
XX CC expression or activity of the polynucleotide and/or polypeptide; and 767
XX CC contig sequences corresponding to the cDNA sequences of the invention
XX CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
XX CC -ADC33394). The nucleic acids and polypeptides of the invention are
XX CC useful in diagnostics, drug screening, forensics, gene mapping, in the
XX CC identification of mutations responsible for genetic disorders or other
XX CC traits, for assessing biodiversity, and in producing many other types of
XX CC data and products dependent on DNA and amino acid sequences. They are
XX CC also used for treating diseases such as Parkinson's disease, Alzheimer's
XX CC disease and other neurodegenerative diseases, anaemia, platelet
XX CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX CC cancer. The nucleic acids may also be used as hybridisation probes or
XX CC primers, and in the recombinant production of a protein. The polypeptides
XX CC are also useful in generating antibodies, as molecular weight markers,
XX CC and as food supplements. The present sequence represents a specifically
XX CC claimed human polypeptide sequence of the invention. Note: The sequence
XX CC data for this patent did not form part of the printed specification, but
XX CC was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 628 AA;

Query Match 99.0%; Score 191; DB 7; Length 628;
Best Local Similarity 97.6%; Pred. No. 4e-13;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EREKQWREKEELMLRLQDXEETKKAERLSEQIQRALQ 41
DB 351 EREKQWREKEELMLRLQDYEEKTKAERLSEQIQRALQ 391

RESULT 8
AAB53356
ID AAB53356 standard; protein; 635 AA.
XX AC AAB53356;
XX DT 09-MAR-2001 (first entry)
XX DE Human colon cancer antigen protein sequence SEQ ID NO:896.
XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX KW identification; cytostatic; cardioactive; neuroprotective; vulnary;
XX KW immunomodulatory; muscular; gynaecological; gastrointestinal;
XX KW nephrotropic; antinefactive; antibacterial; gene therapy; wound;
XX KW neural disorder; immune system disorder; muscular disorder;
XX KW reproductive disorder; gastrointestinal disorder; renal disorder;
XX KW infectious disease; cardiovascular disorder.
XX OS Homo sapiens.
XX PN WO20005351-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US005883.
XX FF
XX XX

PR 12-MAR-1999; 99US-0124270P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX WPI; 2000-587534/55.
DR N-PSDB; AAC98113.
XX Colon cancer associated gene sequences, referred to as colon cancer
XX PT antigens, useful for the treatment, prevention, and diagnosis of colon
XX PT disorders such as colon cancer.
XX PS Claim 11; Page 1449-1451; 2104pp; English.
XX CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
XX CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
XX CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
XX CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and
XX CC vulnary, nephrotropic, antinefactive and antibacterial activities, and
XX CC can be used in gene therapy. The colon cancer antigen polynucleotides,
XX CC proteins and antibodies to the proteins are useful for the prevention,
XX CC treatment and diagnosis of colon disorders such as colon cancer. The
XX CC polynucleotides may be used in diagnostics and research, such as for
XX CC chromosome identification, and as hybridisation probes. The proteins may
XX CC also be used to prevent diseases such as neural disorders, immune system
XX CC disorders, muscular disorders, reproductive disorders, gastrointestinal
XX CC disorders, wounds, renal disorders, infectious diseases, and
XX CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
XX CC sequences used in the exemplification of the present invention
XX SQ Sequence 635 AA;

Query Match 99.0%; Score 191; DB 3; Length 635;
Best Local Similarity 97.6%; Pred. No. 4.1e-13;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EREKQWREKEELMLRLQDXEETKKAERLSEQIQRALQ 41
DB 383 EREKQWREKEELMLRLQDYEEKTKAERLSEQIQRALQ 423

RESULT 9
AAB82020
ID AAB82020 standard; peptide; 34 AA.
XX AC AAB82020;
XX DT 13-JUN-2001 (first entry)
XX DE Human hepreceptor domain B.
XX KW Human; hepreceptor domain B; cytostatic; anti-HIV; antibiotic; nootropic;
XX KW immune response inducer; ezrin; infectious diseases; cancer;
XX KW HIV-related dementia.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 14
XX FT /note= "Optionally phosphorylated"
XX PN GB2354241-A.
XX PD 21-MAR-2001.
XX PF 17-SEP-1999; 99GB-00021881.
XX PR 17-SEP-1999; 99GB-00021881.
XX PA (HOLM/) HOLMS R D.
XX PI Holms RD;
```

XX WPI; 2001-293287/31.
 XX Novel regulatory or unfolding peptides of ezrin that binds to
 PT Heparin, useful for inducing immune response for treating infectious
 PT diseases and cancer.
 XX Claim 5; Page 36; 42pp; English.
 XX The present sequence is domain B of human heparin. The
 CC heparin is a novel active site in human ezrin. Ezrin regulates the
 CC structure of the cortical cytoskeleton to control cell surface
 CC topography. The present invention relates to peptides (see AAB82021 to
 CC AAB82041) that bind to heparin with greater affinity than HEP1 (see
 CC AAB82046). The heparin binding peptides are useful for inducing
 CC immune response, and for treating infectious diseases, cancer and HIV-
 CC related dementia. The present sequence assembles into two anti-parallel
 CC helices with heparin domain A (see AAB82019)
 XX Sequence 34 AA;
 SQ
 Query Match 80.8%; Score 156; DB 4; Length 34;
 Best Local Similarity 97.1%; Pred. No. 1.4e-10;
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 8 MREKEELMLRLQDYEEKTKKARELSEQIQALQ 41
 DB 1 MREKEELMLRLQDYEEKTKKARELSEQIQALQ 34
 RESULT 10
 ID AAU33060 standard; protein; 52 AA.
 XX AAU33060;
 XX 18-DEC-2001 (first entry)
 DE Novel human secreted protein #3551.
 XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX Homo sapiens.
 OS WO200179449-A2.
 PN 25-OCT-2001.
 PD 16-APR-2001; 2001WO-US008656.
 PF 18-APR-2000; 2000US-0052929.
 PR 26-JAN-2001; 2001US-00770160.
 DR (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-611725/70.
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX Claim 20; Page 702; 765pp; English.
 XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX Sequence 52 AA;
 SQ
 Query Match 71.5%; Score 138; DB 4; Length 52;
 Best Local Similarity 78.4%; Pred. No. 2.2e-08;
 Matches 29; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 5 EQMREKEELMLRLQDYEEKTKKARELSEQIQALQ 41
 DB 6 EQMREKEELMLRLQDYEEKTKKARELSEQIKREMK 42
 RESULT 11
 ID ADE54611 standard; protein; 576 AA.
 XX ADE54611;
 XX 29-JAN-2004 (first entry)
 DT Human Protein P26038, SEQ ID NO 416.
 DE Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX Homo sapiens.
 OS WO2003016475-A2.
 PN 27-FEB-2003.
 PD 14-AUG-2002; 2002WO-US025765.
 PF 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PA Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 DR GENBANK; P26038.
 DR New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX Claim 1; Page; 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method. an array, a method for identifying an agent

PS Disclosure; Page 107-108; 126pp; English.

PS Disclosure; Page 107-108; 126pp; English.

CC The invention relates to a novel method for identifying a compound that
 CC modulates T lymphocyte activation. The method comprises contacting a T
 CC cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound,
 CC where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic
 CC acid that hybridises to a nucleic acid encoding a polypeptide having a
 CC sequence selected from two 606-amino acid sequence and a 415-amino acid
 CC sequence given in the specification. The method of the invention has
 CC immunosuppressive, antiasthmatic, antiallergic, and antiinflammatory
 CC activity. The method is useful for identifying compounds that modulate
 CC lymphocyte activation and migration, and for monitoring changes in cell
 CC surface marker expression, cytokine production, antibody production,
 CC proliferation and differentiation, and apoptosis, using either cell lines
 CC or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as
 CC drug targets for compounds that suppress or activate lymphocyte
 CC activation and migration, e.g. for the treatment of diseases in which
 CC modulation of the immune response is desired such as delayed type
 CC hypersensitivity reactions, asthma, allergies, graft versus host disease,
 CC and acute and chronic inflammation. Modulators of lymphocyte activation
 CC are useful for treating disorders related T and B cell activation and
 CC migration. The present sequence is used in the exemplification of the
 CC invention.
 XX
 SQ Sequence 577 AA;

Query Match 66.8%; Score 129; DB 6; Length 577;
 Best Local Similarity 63.4%; Pred. No. 3e-06;
 Matches 26; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 1 EREKEQMMREKELMLRLQDXEKTTKKABRELSEQIQRALQ 41
 Db 334 EREKEKIEREKELMRLKQIEQTKKAOQLEEQTRRALE 374

RESULT 14
 ADB70362
 ID ADB70362 standard; protein; 577 AA.

XX ADB70362;
 XX
 XX
 DT 04-DEC-2003 (first entry)
 DE Moesin SEQ ID NO:54.

XX cancer; malignant pleural mesothelioma; WPM; lung adenocarcinoma;
 KW squamous carcinoma; medulloblastoma; prostate cancer; breast cancer;
 KW diffuse large B-cell lymphoma; follicular lymphoma; ovarian cancer;
 KW human.

XX Homo sapiens.

XX WO2003021229-A2.

XX 13-MAR-2003.

XX 05-SEP-2002; 2002WO-US028203.

XX 05-SEP-2001; 2001US-0317389P.

XX 30-AUG-2002; 2002US-00236031.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX Gordon GU, Jensen RV, Gullans SR, Bueno R;

XX WPI; 2003-290233/28.

XX N-PSDB; ADB70361.

XX Diagnosing cancer cells in tissue sample, or determining prognosis or
 PT outcome of cancer patient, by calculating ratio of expression levels of
 PT genes that are differentially expressed in cancer and non cancer tissues.

XX Claim 77; Page 263-264; 396pp; English.

XX The present invention describes a method (M1) for diagnosing the presence

CC of cancer cells or non-cancer cells in a tissue sample, or determining
 CC the prognosis or outcome of a cancer patient. M1 involves providing a set
 CC of genes that are differentially expressed in cancerous or non-cancerous
 CC conditions, determining the expression levels of the set of genes and
 CC calculating a ratio of the expression levels of the differentially
 CC expressed genes. M1 is useful for diagnosing the presence of cancer cells
 CC or non-cancer cells in a tissue sample, where the cancer is malignant
 CC pleural mesothelioma (WPM), lung adenocarcinoma, squamous carcinoma,
 CC medulloblastoma, prostate cancer, breast cancer, diffuse large B-cell
 CC lymphoma, follicular lymphoma and ovarian cancer, and for determining
 CC prognosis or outcome of a cancer patient. The ratio of expression levels
 CC of differentially expressed genes is used as an indicator of cancer type,
 CC cancer class, and/or cancer prognosis, all of which are useful for
 CC determining a course of treatment of a patient. The present sequence
 CC represents a human protein which is used in an example from the present
 CC invention.

XX Sequence 577 AA;

Query Match 66.8%; Score 129; DB 7; Length 577;
 Best Local Similarity 63.4%; Pred. No. 3e-06;
 Matches 26; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 1 EREKEQMMREKELMLRLQDXEKTTKKABRELSEQIQRALQ 41
 Db 334 EREKEKIEREKELMRLKQIEQTKKAOQLEEQTRRALE 374

RESULT 15
 ADC77532

ID ADC77532 standard; protein; 577 AA.

XX ADC77532;

XX 01-JAN-2004 (first entry)

XX Human moesin protein SEQ ID NO:27.

XX hybridisation; testicular tumour differentially expressed gene 1; TDE1;
 KW patchy vessel; PTV; HLA class II region expressed gene 4; HKE4;
 KW vascular endothelial growth factor; VEGF; antirheumatic; antiarthritic;
 KW antiarteriosclerotic; antidiabetic; ophthalmological; antipsoriatic;
 KW cycostatic; osteopathic; antiangiogenic; antisense gene therapy;
 KW aberrant vascular activity disorder; angiogenesis; vasculogenesis;
 KW cartilage formation disorder; bone formation disorder;
 KW rheumatoid arthritis; atherosclerosis; diabetes; retinopathies;
 KW psoriasis; cancer; human; moesin.

XX Homo sapiens.

XX WO2003066829-A2.

XX 14-AUG-2003.

XX 07-FEB-2003; 2003WO-US003747.

XX 07-FEB-2002; 2002US-0354978P.

XX (DISC-) DISCOVERY GENOMICS INC.

XX Hackett PB, Nasevicius A, Wadman S, Essner J, Larson J, Clark KJ;

XX Roberg-Perez S, Ekker SC;

XX WPI; 2003-731500/69.

XX N-PSDB; ADC77530, ADC77531.

XX New nucleic acid useful for diagnosing, preventing or treating diseases
 PT associated with aberrant vascular activity, angiogenesis, vasculogenesis
 PT and cartilage or bone formation, e.g. atherosclerosis or diabetes.

XX Claim 25; SEQ ID NO 27; 70pp; English.

XX The present invention describes an isolated nucleic acid (I) comprising:

Sequence 577 AA:

Qy	1	E	R	E	K	E	Q	M	R	E	K	E	L	M	L	R	L	Q	D	X	E	E	K	T	K	A	E	R	S	E	Q	I	R	A	L	Q	41
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Dp	334	E	K	E	K	I	E	R	E	K	E	L	M	R	L	K	O	I	E	O	T	T	K	A	O	O	E	L	E	O	T	R	A	L	E	374	

Search completed: August 16, 2004, 09:43:46
Job time : 100 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:45:46 ; Search time 19 Seconds
(without alignments)
111.403 Million cell updates/sec

Title: US-09-856-070b-29
Perfect score: 193
Sequence: 1 EREKQMMREKELMLRLQD.....EETKKAERLSEQIQALQ 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCUS COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	99.0	586	4	US-09-040-725A-1 Sequence 1, Appl
2	73	37.8	1898	1	US-08-056-200-94 Sequence 94, Appl
3	73	37.8	1898	2	US-08-800-644-94 Sequence 94, Appl
4	68	35.2	1180	4	US-09-543-681A-6436 Sequence 6436, Ap
5	67.5	35.0	64	4	US-09-205-258-580 Sequence 580, App
6	67.5	35.0	567	4	US-09-205-258-573 Sequence 573, App
7	66	34.2	764	1	US-08-375-300-4 Sequence 4, Appl
8	66	34.2	764	3	US-09-177-431-4 Sequence 4, Appl
9	66	34.2	764	5	PCT-US95-16930-4 Sequence 2, Appl
10	66	34.2	1089	1	US-08-375-300-2 Sequence 2, Appl
11	66	34.2	1089	3	US-09-177-431-2 Sequence 2, Appl
12	66	34.2	1089	5	PCT-US95-16930-2 Sequence 2, Appl
13	65.5	33.9	740	1	US-08-257-073-5 Sequence 5, Appl
14	65.5	33.9	818	4	US-09-134-000C-6355 Sequence 6355, Ap
15	65.5	33.9	1162	2	US-08-728-323A-2 Sequence 2, Appl
16	65.5	33.9	1162	4	US-09-298-568-2 Sequence 2, Appl
17	65.5	33.9	1162	4	US-09-410-399-2 Sequence 2, Appl
18	64	33.2	2101	1	US-08-466-390-4 Sequence 4, Appl
19	64	33.2	2101	1	US-08-470-950-4 Sequence 4, Appl
20	64	33.2	2101	1	US-08-467-781-4 Sequence 4, Appl
21	64	33.2	2101	1	US-08-195-487-4 Sequence 4, Appl
22	64	33.2	2101	2	US-08-483-924-4 Sequence 4, Appl
23	64	33.2	2101	3	US-09-452-294-1 Sequence 1, Appl
24	64	33.2	2101	5	PCT-US93-06160-4 Sequence 4, Appl
25	62.5	32.4	1326	4	US-09-688-188B-15 Sequence 15, Appl
26	62.5	32.4	1326	4	US-09-291-417D-15 Sequence 15, Appl
27	62	32.1	475	4	US-09-370-838-193 Sequence 193, App

ALIGNMENTS

RESULT 1
US-09-040-725A-1
; Sequence 1, Application US/09040725A

; Patent No. 6399584
; GENERAL INFORMATION:

; APPLICANT: Institut Curie

; APPLICANT: CNRS

; APPLICANT: Arpin, Monique

; APPLICANT: Crepaldi, Tiziana

; APPLICANT: Gauthreau, Alexis

; APPLICANT: Louvard, Daniel

; TITLE OF INVENTION: Pharmaceutical composition containing ezrin mutated

; TITLE OF INVENTION: on tyrosine 353

; FILE REFERENCE: 391082000100

; CURRENT APPLICATION NUMBER: US/09/040,725A

; CURRENT FILING DATE: 1998-03-18

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 586

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-040-725A-1

Query Match 99.0%; Score 191; DB 4; Length 586;

Best Local Similarity 97.6%; Pred. No. 5.7e-14;

Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EREKQMMREKELMLRLQDYEETKKAERLSEQIQALQ 41

Db 334 EREKQMMREKELMLRLQDYEETKKAERLSEQIQALQ 374

RESULT 2

US-08-056-200-94

; Sequence 94, Application US/08056200

; Patent No. 5616500

; GENERAL INFORMATION:

; APPLICANT: Steinert, Peter M.

; APPLICANT: Lee, Seung-Chul

; APPLICANT: Kim, In-Gyu

; APPLICANT: Chung, Soo-Il

; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and

; TITLE OF INVENTION: Methods of Using Same

; NUMBER OF SEQUENCES: 117

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive, Sixteenth Floor

; CITY: Newport Beach

```

; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056,200
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1898 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-056-200-94

Query Match 37.8%; Score 73; DB 1; Length 1898;
Best Local Similarity 41.2%; Pred. No. 3.2;
Matches 14; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 2 REKEQMMREKEELMLRLQDXEETKKAERLSQ 35
DB 1625 REEQLQEGEEQQLRQERDRKFRREEQQLRRQ 1658

RESULT 3
US-08-800-644-94
; Sequence 94, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichoyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/056,200
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1898 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-800-644-94

Query Match 37.8%; Score 73; DB 2; Length 1898;
Best Local Similarity 41.2%; Pred. No. 3.2;
Matches 14; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 2 REKEQMMREKEELMLRLQDXEETKKAERLSQ 35
DB 1625 REEQLQEGEEQQLRQERDRKFRREEQQLRRQ 1658

RESULT 4
US-09-543-681A-6436
; Sequence 6436, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-05
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6436
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-6436

Query Match 35.2%; Score 68; DB 4; Length 1180;
Best Local Similarity 43.2%; Pred. No. 7;
Matches 16; Conservative 9; Mismatches 10; Indels 2; Gaps 1;

QY 1 BREKEQMMREKEELMLRLQDXEETK--KAERLSQ 35
DB 627 EOEKAKQEKAEQERLARLQAEQEKAKQEKAEQ 663

RESULT 5
US-09-205-258-580
; Sequence 580, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876

```


EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 580
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
US-09-205-258-580

Query Match 35.0%; Score 67.5; DB 4; Length 64;
Best Local Similarity 39.5%; Pred. No. 0.38; Indels 3; Gaps 1;
Matches 15; Conservative 10; Mismatches 10; Indels 3; Gaps 1;
Qy 1 ERKEQMMREKEELMLRLQDXEKTKKARELSEQIQR 38
Db 16 ERAKEREKREKEQ---EEEEKEREKEAERERNRQLER 50
RESULT 6
US-09-205-258-573
; Sequence 573, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 573
LENGTH: 567
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (409)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-573

Query Match 35.0%; Score 67.5; DB 4; Length 567;
Best Local Similarity 39.5%; Pred. No. 3.7;
Matches 15; Conservative 10; Mismatches 10; Indels 3; Gaps 1;

Oy 1 EREKQWKEELMLRLQDXEETKKAERELSEQIOR 38
Db 470 ERAKEREKKEQ---EEEQEREKKAERENRQLER 504

RESULT 7
US-08-375-300-4
Sequence 4, Application US/08375300
Patent No. 5679566
GENERAL INFORMATION:
APPLICANT: Feng, He
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN
TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:

NAME: Fasse, J. P.
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04020/046001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-375-300-4
Query Match 34.2%; Score 66; DB 1; Length 764;
Best Local Similarity 44.4%; Pred. No. 7.4; Indels 12; Gaps 1;
Matches 16; Conservative 6; Mismatches 2; Gaps 1;
Oy 6 QMREKEELMLRLQDXEETKKAERELSEQIORALQ 41
Db 625 EMKRYEYERKLDKEE--RKAEEELERQFQKMQ 658
RESULT 8
US-09-177-431-4
Sequence 4, Application US/09177431
Patent No. 6071700
GENERAL INFORMATION:
APPLICANT: He, Feng
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq For Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/177,431
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/955,472
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/050001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-9806
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-177-431-4
Query Match 34.2%; Score 66; DB 3; Length 764;
Best Local Similarity 44.4%; Pred. No. 7.4;
Matches 16; Conservative 6; Mismatches 12; Indels 2; Gaps 1;

QY 6 OMREKEELMLRLQDXEETKKAERLSEIQIQRALQ 41
Db 625 EMKRYEYERKLDKEE--RKAEELERQFQWMQ 658

RESULT 9
PCT-US95-16930-4
; Sequence 4, Application PC/TUS9516930
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE
; TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF
; TITLE OF INVENTION: NONSENSE-MEDIATED mRNA DECAY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16930
; FILING DATE: 27-DEC-1995
; PRIOR APPLICATION DATA: US 08/375,300
; APPLICATION NUMBER: US 08/375,300
; FILING DATE: 20-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04020/046W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)542-5070
; TELEFAX: (617)542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-16930-4

Query Match 34.2%; Score 66; DB 5; Length 764;
Best Local Similarity 44.4%; Pred. No. 7.4; Indels 12; Gaps 1;
Matches 16; Conservative 6; Mismatches 12; Indels 12; Gaps 1;

QY 6 OMREKEELMLRLQDXEETKKAERLSEIQIQRALQ 41
Db 625 EMKRYEYERKLDKEE--RKAEELERQFQWMQ 658

RESULT 10
US-08-375-300-2
; Sequence 2, Application US/08375300
; Patent No. 5679566
; GENERAL INFORMATION:
; APPLICANT: Fasse, He
; APPLICANT: Jacobson, Allan S.
; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN
; TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street Suite 3100
; CITY: Boston

STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. P.
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04020/046001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-375-300-2

Query Match 34.2%; Score 66; DB 1; Length 1089;
Best Local Similarity 44.4%; Pred. No. 11;
Matches 16; Conservative 6; Mismatches 12; Indels 12; Gaps 1;

QY 6 OMREKEELMLRLQDXEETKKAERLSEIQIQRALQ 41
Db 950 EMKRYEYERKLDKEE--RKAEELERQFQWMQ 983

RESULT 11
US-09-177-431-2
; Sequence 2, Application US/09177431
; Patent No. 6071700
; GENERAL INFORMATION:
; APPLICANT: He, Feng
; APPLICANT: Jacobson, Allan S.
; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
; TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Fast-Seq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/177,431
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/955,472
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07917/050001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-9806
; TELEX: 200154

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-177-431-2
Query Match 34.2%; Score 66; DB 3; Length 1089;
Best Local Similarity 44.4%; Pred. No. 11;
Matches 16; Conservative 6; Mismatches 12; Indels 2; Gaps 1;
QY 6 QMREKEELMLRLQDXEETKKAERELSEQIALQ 41
Db 950 EMKRYEYERKLBDEE--RKAEEELERQFQMMQ 983

RESULT 12
PCI-US95-16930-2
; Sequence 2, Application PC/TUS9516930
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE
; TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF
; TITLE OF INVENTION: NONSENSE-MEDIATED mRNA DECAY
; TITLE OF INVENTION: FUNCTION
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCI-US95/16930
; FILING DATE: 27-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/375,300
; FILING DATE: 20-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04020/046W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)542-5070
; TELEFAX: (617)542-8905
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCI-US95-16930-2
Query Match 34.2%; Score 66; DB 5; Length 1089;
Best Local Similarity 44.4%; Pred. No. 11;
Matches 16; Conservative 6; Mismatches 12; Indels 2; Gaps 1;
QY 6 QMREKEELMLRLQDXEETKKAERELSEQIALQ 41
Db 950 EMKRYEYERKLBDEE--RKAEEELERQFQMMQ 983

RESULT 13
US-08-257-073-5
; Sequence 5, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: de Taisne, Charles
; APPLICANT: Tine, John A.
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 740 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-257-073-5
Query Match 33.9%; Score 65.5; DB 1; Length 740;
Best Local Similarity 42.9%; Pred. No. 8.2;
Matches 15; Conservative 9; Mismatches 8; Indels 3; Gaps 1;
QY 1 ERKEQMREKEELMLRLQDXEETKKAERELSEQ 35
Db 682 EKEKEKEKEKEE---KEKEKEKEKEKEKEE 713

RESULT 14
US-09-134-000C-6355
; Sequence 6355, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
```

Search completed: August 16, 2004, 09:46:37
Job time : 20 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:49:27 ; Search time 46 Seconds
(without alignments)
279.805 Million cell updates/sec

Title: US-09-856-070B-29
Perfect score: 193
Sequence: 1 BREKEQMRKEELMLRLQD.....BEKTKAERLSQIQALQ 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTU5_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	191	99.0	436	14	US-10-106-698-4728
2	191	99.0	586	15	Sequence 4728, Ap
3	191	99.0	586	16	Sequence 117, App
4	191	99.0	630	16	US-10-116-275-117
5	191	99.0	635	9	US-10-408-765A-2372
6	191	99.0	635	9	US-10-408-765A-1160
7	123	66.8	577	15	Sequence 896, App
8	129	66.8	577	15	Sequence 896, App
9	129	66.8	577	16	Sequence 54, Appl
10	129	66.8	577	16	Sequence 27, Appl
11	129	66.8	577	16	Sequence 453, App
12	120	62.2	583	14	US-10-408-765A-453
13	117	60.6	579	15	US-10-648-593-187
14	117	60.6	579	15	Sequence 151, App
15	97.5	50.5	213	14	US-10-205-219-151
					Sequence 21, Appl
					Sequence 24, Appl
					Sequence 32544, A

ALIGNMENTS

RESULT 1

US-10-106-698-4728
; Sequence 4728, Application US/101066698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106.698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patent in Ver. 3.0
; SEQ ID NO 4728
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (382)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4728

Query Match 99.0%; Score 191; DB 14; Length 436;

Best Local Similarity 97.6%; Pred. No. 3.7e-12;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 BREKEQMRKEELMLRLQDXXEETKKAERLSQIQALQ 41
DB 184 BREKEQMRKEELMLRLQDYEEKTKKAERLSQIQALQ 224

```

; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1160
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-1160

Query Match          99.0%; Score 191; DB 16; Length 630;
Best Local Similarity 97.6%; Pred. No. 5.4e-12;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EREKQWREKEELMLRLQDYEEKTKKARELSEQIQRALQ 41
Db 378 EREKQWREKEELMLRLQDYEEKTKKARELSEQIQRALQ 418

RESULT 5
US-09-925-299-896
; Sequence 896, Application US/09925299
; Patent No. US2002005627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1356
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 896
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-299-896

Query Match          99.0%; Score 191; DB 9; Length 635;
Best Local Similarity 97.6%; Pred. No. 5.4e-12;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EREKQWREKEELMLRLQDYEEKTKKARELSEQIQRALQ 41
Db 383 EREKQWREKEELMLRLQDYEEKTKKARELSEQIQRALQ 423

RESULT 6
US-09-925-299-896
; Sequence 896, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08

; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 117
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-116-275-117

Query Match          99.0%; Score 191; DB 15; Length 586;
Best Local Similarity 97.6%; Pred. No. 5e-12;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EREKQWREKEELMLRLQDYEEKTKKARELSEQIQRALQ 41
Db 334 EREKQWREKEELMLRLQDYEEKTKKARELSEQIQRALQ 374

RESULT 3
US-10-408-765A-2372
; Sequence 2372, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2372
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-2372

Query Match          99.0%; Score 191; DB 16; Length 586;
Best Local Similarity 97.6%; Pred. No. 5e-12;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EREKQWREKEELMLRLQDYEEKTKKARELSEQIQRALQ 41
Db 334 EREKQWREKEELMLRLQDYEEKTKKARELSEQIQRALQ 374

RESULT 4
US-10-408-765A-1160
; Sequence 1160, Application US/10408765A
; Publication No. US20040101874A1
```



```

RESULT 10
US-10-408-765A-454
; Sequence 454, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Faby, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 454
; LENGTH: 577

```


Search completed: August 16, 2004, 09:53:27
Job time : 46 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:44:16 ; Search time 16 Seconds
(without alignments)
246.491 Million cell updates/sec

Title: US-09-856-070b-29

Perfect score: 193

Sequence: 1 EREKQWREKSELMLRLQD.....ERTKKARELSEQIQRALQ 41

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	191	99.0	586	1 A34400	ezrin [validated]
2	191	99.0	630	2 T47177	hypothetical prote
3	178	92.2	581	2 I45989	ezrin - bovine
4	173	89.6	586	1 B41129	ezrin - mouse
5	129	66.8	577	1 A41289	moesin - human
6	127	65.8	577	1 S39804	moesin - pig
7	121	62.7	583	1 S39805	radixin - pig
8	120	62.2	583	1 A41129	radixin - mouse
9	119	61.7	583	1 A46127	radixin - human
10	79	40.9	1407	1 S28589	trichohyalin - rab
11	77	39.9	4574	2 G02220	plectin - human
12	75	38.9	1549	1 A40591	trichohyalin - she
13	73	37.8	1898	1 A45973	trichohyalin - hum
14	72	37.3	463	2 B72500	probable seryl-trn
15	71.5	37.0	4687	1 A39638	plectin - rat
16	71	36.8	1069	2 T00377	KIAA0642 protein -
17	71	36.8	1132	2 T43483	translation initia
18	71	36.8	4684	2 A59404	plectin [imported]
19	69.5	36.0	899	2 I50590	class II INCENP pr
20	69.5	36.0	877	2 I50591	class II INCENP pr
21	69.5	36.0	2962	2 T19756	hypothetical prote
22	69	35.8	1039	2 T15985	hypothetical prote
23	68.5	35.5	174	2 E69038	heat shock protein
24	68.5	35.5	226	2 T46368	hypothetical prote
25	68.5	35.5	899	2 D96594	unknown protein, 7
26	68	35.2	743	2 A29232	101k malaria antig
27	68	35.2	1164	2 T24806	hypothetical prote
28	68	35.2	1378	2 G88637	protein F53H1.4 li
29	67.5	35.0	279	2 D71453	hypothetical prote

30	67.5	35.0	1280	2 T00365	hypothetical prote
31	67	34.7	699	2 T21379	hypothetical prote
32	66.5	34.5	400	2 E70318	hypothetical prote
33	66.5	34.5	527	2 S33068	myosin heavy chain
34	66.5	34.5	919	2 T15087	hypothetical prote
35	66.5	34.5	1940	2 A59287	myosin heavy chain
36	66	34.2	175	2 T07210	H+-transporting tw
37	66	34.2	281	2 F75216	hypothetical prote
38	66	34.2	1089	2 S48244	NMD2 protein - yea
39	66	34.2	1231	2 A54803	microtubule-associ
40	66	34.2	1252	2 T00263	hypothetical prote
41	66	34.2	3187	2 JCS837	36k Golgi complex
42	66	34.2	5327	2 T13564	microtubule-associ
43	65.5	33.9	365	2 T28184	hypothetical prote
44	65.5	33.9	756	2 T00367	hypothetical prote
45	65.5	33.9	777	2 T21048	hypothetical prote

ALIGNMENTS

RESULT 1

A34400

ezrin [validated] - human

N:Alternate names: cytoovillin; p81 protein; villin 2

C:Species: Homo sapiens (man)

C>Date: 22-Jun-1990 #sequence_revision 14-Jul-1994 #text_change 24-Nov-2003

C:Accession: A34400; S09263; E61002

R:Turunen, O.; Wingqvist, R.; Pakkanen, R.; Grzeschik, K.H.; Wahlstroem, T.; Vaheri, A.

J. Biol. Chem. 264, 16727-16732, 1989

A:Title: Cytoovillin, a microvillar M-r 75,000 protein. cDNA sequence, prokaryotic express

A:Reference number: A34400; MUID:89380299; PMID:2674140

A:Accession: A34400

A:Molecule type: mRNA

A:Residues: 1-586 <TUR>

A:Cross-references: GB:J05021

A:Note: The translation of residues 1-11 is not given

A:Note: parts of this sequence were confirmed by protein sequencing

R:Gould K.L.; Bretscher A.; Esch, F.S.; Hunter, T.

EMBO J. 8, 4133-4142, 1989

A:Title: cDNA cloning and sequencing of the protein-tyrosine kinase substrate, ezrin, rev

A:Reference number: S09263; MUID:90076135; PMID:2591371

A:Accession: S09263

A:Molecule type: mRNA

A:Residues: 2-586 <GOU>

A:Cross-references: GB:X51521; MID:g31282; PIDN:CRA35893.1; PID:g31283

R:Bauw, G.; Rasmussen, H.H.; Van Den Bulcke, W.; Van Damme, J.; Fuype, M.; Gesser, B.; Ce

Electrophoresis 11, 528-536, 1990

A:Title: Two-dimensional gel electrophoresis, protein electrophoresis and microsequencing

A:Reference number: A61002; MUID:91031404; PMID:1699755

A:Accession: E61002

A:Molecule type: protein

A:Residues: 255-263/194, 'Q', 196-199, 'X', 201-264-270 <BAU>

A:Note: It is not certain whether this material represents ezrin or radixin (see entry A:

A:Note: this material corresponds to transformed epithelial amnion cell (AWA) database p:

C:Comment: This protein is located in microvilli and is proposed to play a role in modul:

C:Genetics:

A:Gene: GDB:VIL2

A:Cross-references: GDB:L20489; OMIM:123900

A:Map position: 6225-6q26

C:Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology

C:Keywords: actin binding; cytoskeleton; membrane-associated protein; phosphoprotein

F:2-586/Product: ezrin #status experimental <MAP>

F:7-291/Domain: protein 4.1 membrane-binding domain homology <B41>

F:553-586/Region: actin binding #status predicted

F:66/Binding site: phosphate (Ser) (covalent) #status predicted

F:214,299,332/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 99.0%; Score 191; DB 1; Length 586;

Best Local Similarity 97.6%; Pred. No. 5.9e-11;

Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EREKQWREKSELMLRLQDXXEKKARELSEQIQRALQ 41

```

Db      334 EREKQMRKEBELMLRLQDYEEKTKKAERLSEIQIALQ 374
|||||
RESULT 2
T47177
Hypothetical protein DKFp762H157.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 24-Nov-2003
C:Accession: T47177
R:Otterwaelder, B.; Obermaier, B.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24377
A:Accession: T47177
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-630 <AAA>
A:Cross-references: EMBL:AL462086
A:Experimental source: adult melanoma (MeWo cell line); clone DKFp762H157
C:Genetics:
A>Note: DKFp762H157.1
C:Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology
Query Match 99.0%; Score 191; DB 2; Length 630;
Best Local Similarity 97.6%; Pred. No. 6.3e-11;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EREKQMRKEBELMLRLQDYEEKTKKAERLSEIQIALQ 41
|||||
Db      378 EREKQMRKEBELMLRLQDYEEKTKKAERLSEIQIALQ 418
|||||
RESULT 3
I45889
ezrin - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 24-Nov-2003
C:Accession: I45889
R:Bergson, C.M.; Zhao, H.; Saijoh, K.; Duman, R.S.; Nestler, E.J.
Mol. Cell. Neurosci. 4, 64-73, 1993
A:Title: Ezrin and osteonectin, two proteins associated with cell shape and growth, are
A:Reference number: I45889
A:Accession: I45889
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-581 <BE>
A:Cross-references: GB:M98493; NID:G289407; PID:AAA30510.1; PID:G289408
C:Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology
F:7-291/Domain: protein 4.1 membrane-binding domain homology <B41>
Query Match 92.2%; Score 178; DB 2; Length 581;
Best Local Similarity 87.8%; Pred. No. 9.8e-10;
Matches 36; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 EREKQMRKEBELMLRLQDYEEKTKKAERLSEIQIALQ 41
|||||
Db      334 EREKQMRKEBELMLRLQDYEEKTKKAERLSEIQIALQ 374
|||||
RESULT 4
B41129
ezrin - mouse
N:Alternate names: cyto villin; p81 protein; radixin; villin 2
C:Species: Mus musculus (house mouse)
C>Date: 03-Aug-1992 #sequence_revision 14-Jul-1994 #text_change 24-Nov-2003
C:Accession: B41129; C46501; A46501; B46501; S24200
R:Funayama, N.; Nagatuchi, A.; Sato, N.; Tsukita, S.; Tsukita, S.
J. Cell Biol. 115, 1039-1048, 1991
A:Title: Radixin is a novel member of the band 4.1 family.
A:Reference number: A41129; MUID:92064635; PMID:1955455
A:Accession: B41129
A:Molecule type: mRNA
A:Residues: 1-586 <FUN>

```

```

A:Cross-references: EMBL:X60671; NID:G50880; PIDN:CAA43086.1; PID:G50881
R:Egerton, M.; Burgess, W.H.; Chen, D.; Drucker, B.J.; Bretscher, A.; Samelson, L.E.
J. Immunol. 149, 1847-1852, 1992
A:Title: Identification of ezrin as an 81-kDa tyrosine-phosphorylated protein in T cells.
A:Reference number: A46501; MUID:92388649; PMID:1381389
A:Accession: C46501
A>Status: preliminary
A:Molecule type: protein
A:Residues: 412-426 <EGE>
A:Experimental source: MRL lpr/lpr, T-cells
A:Note: sequence extracted from NCBI backbone (NCBIP:112938)
A:Accession: A46501
A>Status: preliminary
A:Molecule type: protein
A:Residues: 27-33, 'E' <EG2>
A:Experimental source: MRL lpr/lpr, T-cells
A:Note: sequence extracted from NCBI backbone (NCBIP:112936)
A:Accession: B46501
A>Status: preliminary
A:Molecule type: protein
A:Residues: 53-57, 148, 'L', 150, 'G', 152-155 <EG3>
A:Experimental source: MRL lpr/lpr, T-cells
A:Note: sequence extracted from NCBI backbone (NCBIP:112940)
C:Comment: This protein is located in microvilli and is proposed to play a role in modulate
C:Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology
C:Keywords: actin binding; cytoskeleton; cytosol; membrane-associated protein; phosphoprotein
F:2-586/Product: ezrin #status predicted <MAT>
F:7-291/Domain: protein 4.1 membrane-binding domain homology <B41>
F:553-586/Region: actin binding #status predicted
F:66/Binding site: phosphate (Ser) (covalent) #status predicted
F:214,239,332/Binding site: phosphate (Thr) (covalent) #status predicted
Query Match 89.8%; Score 173; DB 1; Length 586;
Best Local Similarity 82.9%; Pred. No. 2.9e-09;
Matches 34; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
QY 1 EREKQMRKEBELMLRLQDYEEKTKKAERLSEIQIALQ 41
|||||
Db      334 EREKQMRKEBELMLRLQDYEEKTKKAERLSEIQIALQ 374
|||||
RESULT 5
A41289
moesin - human
N:Alternate names: membrane-organizing extension spike protein
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1992 #sequence_revision 14-Jul-1994 #text_change 24-Nov-2003
C:Accession: A41289
R:Lankes, W.T.; Furthmayr, H.
Proc. Natl. Acad. Sci. U.S.A. 88, 8297-8301, 1991
A:Title: Moesin: a member of the protein 4.1-talin-ezrin family of proteins.
A:Reference number: A41289; MUID:92020840; PMID:1924289
A:Accession: A41289
A:Molecule type: mRNA
A:Residues: 1-577 <LAN>
A:Cross-references: GB:M69066; NID:G188625; PIDN:AAA36322.1; PID:G188626
C:Comment: Moesin is proposed to be involved in linking the cytoskeleton to the plasma me
C:Genetics:
A:Gene: GDB:MSN
A:Cross-references: GDB:136819; OMIM:309845
A:Map position: Xq11.2-Xq12
C:Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology
C:Keywords: actin binding; cytoskeleton; membrane protein
F:2-577/Product: moesin #status predicted <MAT>
F:7-291/Domain: protein 4.1 membrane-binding domain homology <B41>
F:544-577/Region: actin binding #status predicted
Query Match 66.8%; Score 129; DB 1; Length 577;
Best Local Similarity 63.4%; Pred. No. 4.1e-05;
Matches 26; Conservative 9; Mismatches 6; Indels 0; Gaps 0;
QY 1 EREKQMRKEBELMLRLQDYEEKTKKAERLSEIQIALQ 41
|||||

```

```

QY      3 EKEQMREKBEMLRLQDXEKKYKAEKELSEQI 36
      |:::|:::|:::|:::|:::|:::|:::|:::|
DB     73 EERLIAEAKLLREREDLEELKKKVERERBEAL 106

RESULT 15
A39638
Plectin - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: A39638; S21876
R:Wiche, G.; Becker, B.; Huber, K.; Weitzer, G.; Castanon, M.J.; Hauptmann, R.; Stratowa
J' Cell Biol. 114, 83-99, 1991
A:Title: Cloning and sequencing of rat plectin indicates a 466-kD polypeptide chain with
A:Reference number: A39638; MUID:91268156; PMID:2050743
A:Accession: A39638
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4687 <WIC>
A:Cross-references: EMBL:X59601; NID:G1292885; PIDN:CAA42169.1; PID:G1561642
C:Superfamily: plectin, alpha-actinin actin-binding domain homology; ribosomal protein S
F;6-103/Domain: cytoskeleton; transmembrane protein
F;184-399/Domain: alpha-actinin actin-binding domain homology <ACT>

Query Match      37.0%; Score 71.5; DB 1; Length 4687;
Best Local Similarity 31.8%; Pred. No. 78;
Matches 14; Conservative 15; Mismatches 12; Indels 3; Gaps 1;

QY      1 EREKQMRKBEMLRLQDXEKKYKAE---RELSEQIQRLQ 41
      |:::|:::|:::|:::|:::|:::|:::|
DB     2675 QRQQQMEQEKQELVASMEARRRQREAEQGVRRKQELQRLQ 2718

```

Search completed: August 16, 2004, 09:46:07
Job time : 17 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:39:51 ; Search time 13 seconds
(without alignments)

164.221 Million cell updates/sec

Title: US-09-856-070B-29

Perfect score: 193

Sequence: 1 BREKEQMMREKEELMLRLQD.....EETKKAERLSEQIQRALQ 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	99.0	585	1 EZRI HUMAN	P15311 homo sapien
2	178	92.2	580	1 EZRI BOVIN	P31976 bos taurus
3	173	89.6	585	1 EZRI MOUSE	P26040 mus musculus
4	173	89.6	585	1 EZRI RABIT	Q8h2q5 oryctolagus
5	129	66.8	576	1 MOES HUMAN	P26038 homo sapien
6	127	65.8	576	1 MOES PIG	P26042 sus scrofa
7	126	65.3	576	1 MOES MOUSE	P26041 mus musculus
8	126	65.3	576	1 MOES RAT	Q35763 rattus norv
9	124	64.2	583	1 RADI CHICK	Q9pu45 gallus gall
10	121	62.7	583	1 RADI PIG	P26044 sus scrofa
11	120	62.2	583	1 RADI MOUSE	P26043 mus musculus
12	119	61.7	583	1 RADI HUMAN	P35241 homo sapien
13	79	40.9	1407	1 TRHY RABIT	P37709 oryctolagus
14	75	38.9	1549	1 TRHY SHEEP	P22793 ovis aries
15	73	37.8	1898	1 TRHY HUMAN	Q07283 homo sapien
16	72	37.3	463	1 SYS AERPE	Q9yag3 aeropyrum p
17	72	37.3	463	1 PLEI CRIGR	Q9ji55 cricetus
18	71.5	37.0	4687	1 PLEI RAT	P30427 rattus norv
19	71	36.8	1220	1 FLEP HUMAN	Q60841 homo sapien
20	71	36.8	1220	1 FLEI HUMAN	Q15149 homo sapien
21	70.5	36.5	1308	1 M4K6 MOUSE	Q9jms2 mus musculus
22	70.5	36.5	1332	1 M4K6 HUMAN	Q844c8 homo sapien
23	69.5	36.0	877	1 INCE CHICK	P53352 gallus gall
24	69.5	36.0	1338	1 ACIN MOUSE	Q9jix8 mus musculus
25	69	35.8	1009	1 YLKE CAEEL	P41954 caenorhabdi
26	68.5	35.5	174	1 GRPE METHF	Q27350 methanobact
27	68	35.2	743	1 ABRA PLAFQ	Q27320 plasmodium
28	67.5	35.0	1341	1 ACIN HUMAN	Q9ukv3 homo sapien
29	66.5	34.5	329	1 IMP4 HUMAN	Q9nuv9 homo sapien
30	66.5	34.5	693	1 GOA6 HUMAN	Q9ny43 homo sapien
31	66	34.2	175	1 ATPF CHLUV	P56296 chlrella v
32	66	34.2	1089	1 NMD2 YEAST	P38798 saccharomyc
33	66	34.2	1231	1 XF43_MOUSE	P33174 mus musculus

ALIGNMENTS

RESULT 1

EZRI_HUMAN STANDARD; Q96CU8; PRT; 585 AA.
AC P15311; P23714; Q96CU8; Q9NSJ4;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB Ezrin (p81) (Cytovillin) (Villin 2).
GN VIL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90076135; PubMed=2591371;
RA Gould K.L., Bretscher A., Esch F.S., Hunter T.;
RT "cDNA cloning and sequencing of the protein-tyrosine kinase
substrate, ezrin, reveals homology to band 4.1.";
RL EMBO J. 8:4133-4142 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta.
RX MEDLINE=89380299; PubMed=2674140;
RA Turunen O., Winqvist R., Pakkanen R., Grzeschik K.-H., Wahlstrom T.,
Vaehri A.;
RT "Cytovillin, a microvillar Mr 75,000 protein. cDNA sequence,
prokaryotic expression, and chromosomal localization.";
RL J. Biol. Chem. 264:16727-16732 (1989).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT LEU-531.
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Weil B., Wiemann S.;
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT LEU-531.
RC TISSUE=Colon.
RX MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., USCIN T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madañ A., Rodriguez S., Sanchez A.,
Whiting M., Madañ A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";

34 66 34.2 5560 1 SPEN DROME
35 65.5 33.9 640 1 GBP4 HUMAN
36 65.5 33.9 756 1 EFER HUMAN
37 65 33.7 854 1 DIS1 HUMAN
38 65 33.7 919 1 INCE HUMAN
39 64.5 33.4 911 1 CAFN_MOUSE
40 64.5 33.4 1023 1 GLT_DROME
41 64.5 33.4 1240 1 YNJI_YEAST
42 64 33.2 958 1 IF3A_TOBAC
43 64 33.2 1085 1 YAFN SCHPO
44 64 33.2 8797 1 SNEI HUMAN
45 63.5 32.9 113 1 PFDE_METJA

Q8sx83 drosophila
Q96pp9 homo sapien
Q75154 homo sapien
Q9nr15 homo sapien
Q9ngs7 homo sapien
Q9qwf0 mus musculus
P33438 drosophila
P53935 saccharomyc
Q40554 nicotiana t
Q09863 schizosacch
Q8nf91 homo sapien
Q58394 methanococc

```

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
SEQUENCE OF 171-179 AND 342-349.
MEDLINE=96311348; PubMed=8713105;
Egerton M., Moritz R.L., Druker B., Kello A., Simpson R.J.;
"Identification of the 70kD heat shock cognate protein (Hsc70) and
alpha-actinin-1 as novel phosphotyrosine-containing proteins in T
lymphocytes.";
Biochem. Biophys. Res. Commun. 224:666-674(1996).
[6]
PHOSPHORYLATION BY PDGFR.
MEDLINE=92406868; PubMed=1382070;
Krieg J., Hunter T.;
"Identification of the two major epidermal growth factor-induced
tyrosine phosphorylation sites in the microvillar core protein
ezrin.";
J. Biol. Chem. 267:19258-19265(1992).
[7]
PHOSPHORYLATION.
MEDLINE=92388649; PubMed=1381389;
Egerton M., Burgess W.H., Chen D., Druker B.J., Bretscher A.,
Samelson L.E.;
"Identification of ezrin as an 81-kDa tyrosine-phosphorylated protein
in T cells.";
J. Immunol. 149:1847-1852(1992).
-!- FUNCTION: Probably involved in connections of major cytoskeletal
structures to the plasma membrane.
-!- SUBCELLULAR LOCATION: Microvillar peripheral membrane protein
(cytoplasmic side).
-!- TISSUE SPECIFICITY: Component of the microvilli of intestinal
epithelial cells.
-!- PTM: Phosphorylated by tyrosine-protein kinases.
-!- SIMILARITY: Contains 1 FERM domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X51521; CA935893.1; -
EMBL; J05021; AA81278.1; ALT_INIT.
EMBL; AL162086; CAB82418.1; ALT_INIT.
EMBL; BC013903; AAH13903.1; -
PIR; A34400; A34400.
PDB; 1NT2; 25-FEB-03.
DR SWISS-2DPAGE; P15311; HUMAN.
DR Genew; HGNC:12691; VIL2.
DR MIM; 123900; -
DR InterPro; IPR000299; Band 4.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR InterPro; IPR008954; Moesin.
DR Pfam; PF00373; Band 41; 1.
DR Pfam; PF00769; ERM; 1.
DR PRINTS; PR00935; BAND41.
DR SMART; SM00295; B41; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS00507; FERM_3; 1.
DR Structural protein; Cytoskeleton; Phosphorylation; Polymorphism;
3D-structure.
FT INIT_MET 0 0
FT DOMAIN 1 294 FERM.
FT MOD_RES 145 145 PHOSPHORYLATION (BY PDGFR).
FT MOD_RES 353 353 PHOSPHORYLATION (BY PDGFR).
FT VARIANT 531 531 V -> L.
FT /FTID=VAR_015112.
SQ SEQUENCE 585 AA; 69267 MW; 2A844D140B3B06CC CRC64;

Query Match 99.0%; Score 191; DB 1; Length 585;
Best Local Similarity 97.8%; Pred. No. 1.1e-10;

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
SEQUENCE OF 171-179 AND 342-349.
MEDLINE=96311348; PubMed=8713105;
Egerton M., Moritz R.L., Druker B., Kello A., Simpson R.J.;
"Identification of the 70kD heat shock cognate protein (Hsc70) and
alpha-actinin-1 as novel phosphotyrosine-containing proteins in T
lymphocytes.";
Biochem. Biophys. Res. Commun. 224:666-674(1996).
[6]
PHOSPHORYLATION BY PDGFR.
MEDLINE=92406868; PubMed=1382070;
Krieg J., Hunter T.;
"Identification of the two major epidermal growth factor-induced
tyrosine phosphorylation sites in the microvillar core protein
ezrin.";
J. Biol. Chem. 267:19258-19265(1992).
[7]
PHOSPHORYLATION.
MEDLINE=92388649; PubMed=1381389;
Egerton M., Burgess W.H., Chen D., Druker B.J., Bretscher A.,
Samelson L.E.;
"Identification of ezrin as an 81-kDa tyrosine-phosphorylated protein
in T cells.";
J. Immunol. 149:1847-1852(1992).
-!- FUNCTION: Probably involved in connections of major cytoskeletal
structures to the plasma membrane.
-!- SUBCELLULAR LOCATION: Microvillar peripheral membrane protein
(cytoplasmic side).
-!- TISSUE SPECIFICITY: Component of the microvilli of intestinal
epithelial cells.
-!- PTM: Phosphorylated by tyrosine-protein kinases.
-!- SIMILARITY: Contains 1 FERM domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X51521; CA935893.1; -
EMBL; J05021; AA81278.1; ALT_INIT.
EMBL; AL162086; CAB82418.1; ALT_INIT.
EMBL; BC013903; AAH13903.1; -
PIR; A34400; A34400.
PDB; 1NT2; 25-FEB-03.
DR SWISS-2DPAGE; P15311; HUMAN.
DR Genew; HGNC:12691; VIL2.
DR MIM; 123900; -
DR InterPro; IPR000299; Band 4.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR InterPro; IPR008954; Moesin.
DR Pfam; PF00373; Band 41; 1.
DR Pfam; PF00769; ERM; 1.
DR PRINTS; PR00935; BAND41.
DR SMART; SM00295; B41; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS00507; FERM_3; 1.
DR Structural protein; Cytoskeleton; Phosphorylation; Polymorphism;
3D-structure.
FT INIT_MET 0 0
FT DOMAIN 1 294 FERM.
FT MOD_RES 145 145 PHOSPHORYLATION (BY PDGFR).
FT MOD_RES 353 353 PHOSPHORYLATION (BY PDGFR).
FT VARIANT 531 531 V -> L.
FT /FTID=VAR_015112.
SQ SEQUENCE 585 AA; 69267 MW; 2A844D140B3B06CC CRC64;

Query Match 99.0%; Score 191; DB 1; Length 585;
Best Local Similarity 97.8%; Pred. No. 1.1e-10;

Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EREKEQWREKELMLRLQDYEKKTKKARELSEQIQALQ 41
Db 333 EREKEQWREKELMLRLQDYEKKTKKARELSEQIQALQ 373
RESULT 2
EZRI_BOVIN STANDARD; PRT; 580 AA.
AC P31976; 1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ezrin (p81) (Cytovillin) (Villin 2).
GN VIL2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Bergson C.M., Zhao H., Saijoh K., Duman R.S., Nestler E.J.;
RT "Ezrin and osteonectin, two proteins associated with cell shape and
RT growth, are enriched in the locus coeruleus.";
RL Mol. Cell. Neurosci. 4:64-73(1993).
RN [2]
RP SEQUENCE OF 1-15 AND 126-140.
RC TISSUE=Kidney;
RA Galat A., Gerbod M.C., Bouet F., Riviere S.;
RT "Proteins and their amino acid compositions: uniqueness, variability,
RT and applications.";
RL Arch. Biochem. Biophys. 330:229-237(1996).
-!- FUNCTION: Probably involved in connections of major cytoskeletal
structures to the plasma membrane.
-!- SUBCELLULAR LOCATION: Microvillar peripheral membrane protein
(cytoplasmic side).
-!- PTM: Phosphorylated by tyrosine-protein kinases.
-!- SIMILARITY: Contains 1 FERM domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M98498; AAA30510.1; -
PIR; I45889; I45889.
DR InterPro; IPR000299; Band 4.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR InterPro; IPR008954; Moesin.
DR Pfam; PF00373; Band 41; 1.
DR Pfam; PF00769; ERM; 1.
DR PRINTS; PR00935; BAND41.
DR SMART; SM00295; B41; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS00507; FERM_3; 1.
DR Structural protein; Cytoskeleton; Phosphorylation.
FT INIT_MET 0 0
FT DOMAIN 1 294 FERM.
FT MOD_RES 145 145 PHOSPHORYLATION (BY PDGFR).
FT MOD_RES 353 353 PHOSPHORYLATION (BY PDGFR).
FT /FTID=VAR_015112.
SQ SEQUENCE 580 AA; 68629 MW; ECD663E5C200FAA3 CRC64;

Query Match 92.2%; Score 178; DB 1; Length 580;
```

```
Best Local Similarity 87.8%; Pred. No. 1.7e-09;
Matches 36; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EREKQMMREKEELMLRLQDYVEKTKKAERLSQIQALQ 41
DB 333 EREKQMMREKEELMLRLQDYVEKTKKAERLSQIQALQ 373

RESULT 3
EZRI_MOUSE
ID EZRI_MOUSE STANDARD; PRT; 585 AA.
AC F26040; Q80ZTS; Q9DC11;
DT 01-MAY-1992 (Rel. 22, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ezrin (p81) (Cytovillin) (Villin 2).
GN VIL2
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92064635; PubMed=1955455;
RA Funayama N., Nagafuchi A., Sato N., Tsukita S.;
RT "Radixin is a novel member of the band 4.1 family.";
RL J. Cell Biol. 115:1039-1048 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Yasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Sato R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RP "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Srausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```

```
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.O.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal
CC structures to the plasma membrane.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Component of the microvilli of intestinal
CC epithelial cells.
CC -!- PTM: Phosphorylated by tyrosine-protein kinases.
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60671; CAA43086.1; --
CC EMBL; AK002766; BAB2341.1; --
CC EMBL; BC048181; AAH48181.2; --
CC PIR; B41129; B41129.
CC MGD; MGI:98931; V112.
CC InterPro; IPR000299; Band_4.1.
CC InterPro; IPR000798; Pz/rad/moesin.
CC Pfam; PF00373; Band_41; 1.
CC Pfam; PF00769; ERM; 1.
CC SMART; SMC0295; BAND41.
CC PRINTS; PR00935; BAND41.
CC PROSITE; PS00660; FERM_1; 1.
CC PROSITE; PS00661; FERM_2; 1.
CC PROSITE; PS00057; FERM_3; 1.
CC Structural protein; Cytoskeleton; Phosphorylation.
FT INIT MET 0 0 BY SIMILARITY.
FT MOD_RES 145 145 FERM.
FT MOD_RES 353 353 PHOSPHORYLATION (BY PDGFR)
FT MOD_RES 353 353 (BY SIMILARITY).
FT MOD_RES 353 353 PHOSPHORYLATION (BY PDGFR)
FT MOD_RES 353 353 (BY SIMILARITY).
FT CONFLICT 47 47 Q -> P (IN REF. 1).
FT CONFLICT 324 324 T -> A (IN REF. 1).
FT CONFLICT 569 569 Q -> R (IN REF. 2).
SQ SEQUENCE 585 AA; 69275 MW; 205CED2E326A0562 CRC64;

Query Match 89.6%; Score 173; DB 1; Length 585;
Best Local Similarity 82.9%; Pred. No. 4.8e-09;
Matches 34; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 EREKQMMREKEELMLRLQDYVEKTKKAERLSQIQALQ 41
DB 333 EREKQMMREKEELMLRLQDYVEKTKKAERLSQIQALQ 373

RESULT 4
EZRI_MOUSE
ID EZRI_MOUSE STANDARD; PRT; 585 AA.
AC Q8HZ05;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ezrin (p81) (Cytovillin) (Villin 2).
GN VIL2
OS Cryptolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


Query Match 66.8%; Score 129; DB 1; Length 576;
 Best Local Similarity 63.4%; Pred. No. 4.8e-05;
 Matches 26; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 EREKQMMREKELMLRLQDYEEKTKKABRELSEIQIALQ 41
 DB 333 EKEKIEREKELMERLKQIEEQTKAQOELEQTRRAL 373

RESULT 6

MOES_PIG STANDARD; PRT; 576 AA.
 AC P26042;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Moesin (Membrane-organizing extension spike protein).
 GN MSN.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94092743; PubMed=8268231;
 RA Lankes W.T., Schwartz-Albiez R., Furtmayr H.;
 RT "Cloning and sequencing of porcine moesin and radixin cDNA and
 RT identification of highly conserved domains."
 RL Biochim. Biophys. Acta 1216:478-482(1993).
 CC !- FUNCTION: Probably involved in connections of major cytoskeletal
 CC structures to the plasma membrane.
 CC !- SIMILARITY: Contains 1 FERM domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M86450; AAB02864.1; -.
 DR PIR; S39804; S39804.
 DR InterPro; IPR000299; Band 4.1.
 DR InterPro; IPR000798; Ez/rad/moesin.
 DR InterPro; IPR008954; Moesin.
 DR Pfam; PF00373; Band 4; 1.
 DR Pfam; PF00769; ERM; 1.
 DR PRINTS; PR00935; BAND41.
 DR SMART; SM00295; B41; 1.
 DR PROSITE; PS00660; FERM 1; 1.
 DR PROSITE; PS00661; FERM 2; 1.
 DR PROSITE; PS00657; FERM 3; 1.
 KW Structural protein; Cytoskeleton.
 FT INIT MET 0 BY SIMILARITY.
 FT DOMAIN 1 294
 SQ SEQUENCE 576 AA; 67529 MW; 3CE0BF78E5F75B6B CRC64;

Query Match 65.8%; Score 127; DB 1; Length 576;
 Best Local Similarity 65.0%; Pred. No. 7.3e-05;
 Matches 26; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 1 EREKQMMREKELMLRLQDYEEKTKKABRELSEIQIALQ 40
 DB 333 EKEKIEREKELMERLKQIEEQTKAQOELEQTRRAL 372

RESULT 7

MOES_MOUSE STANDARD; PRT; 576 AA.
 AC P26041;
 DT 01-MAY-1992 (Rel. 22, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Moesin (Membrane-organizing extension spike protein).
 GN MSN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93055012; PubMed=1429901;
 RA Sato N., Funayama N., Nagafuchi A., Yonemura S., Tsukita S.,
 RA Tsukita S.;
 RT "A gene family consisting of ezrin, radixin and moesin. Its specific
 RT localization at actin filament/plasma membrane association sites."
 RL J. Cell Sci. 103:131-143(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nkaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Ienhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sancelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Breast tumor;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore I., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Boraldo M.F., Casavant T.D., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

RN SEQUENCE OF 11-576 FROM N.A.
RP MEDLINE=92243764; PubMed=1573844;
RA Furthmayr H., Lankes W.T., Anisava M.R.;
RT "Moesin, a new cytoskeletal protein and constituent of filopodia: its
RT role in cellular functions."
RL Kidney Int. 41:665-670(1992).
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal
CC structures to the plasma membrane.
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S47577; AAA11762.1; -
CC EMBL; AK088336; BAC40290.1; -
CC EMBL; BC047366; BAC47366.1; -
CC EMBL; M86390; AAA39728.1; -
CC MGD; MGI:97167; Msn.
CC InterPro; IPR000299; Band 4.1.
CC InterPro; IPR000798; Ez/rad/moesin.
CC Pfam; PF00373; Band 41; 1.
CC Pfam; PF00769; ERM_1.
CC PRINTS; PR00935; BAND41.
CC SMART; SM00295; B41; 1.
CC PROSITE; PS00660; FERM_1; 1.
CC PROSITE; PS00661; FERM_2; 1.
CC PROSITE; PS00507; FERM_3; 1.
CC Structural protein; Cytoskeleton.
CC KW Structural protein; Cytoskeleton.
CC FT INIT MET 0 BY SIMILARITY.
CC DOMAIN 1 294 FERM.
CC SEQUENCE 576 AA; 67635 MW; 5E0F455552E9145 CRC64;

Query Match 65.3%; Score 126; DB 1; Length 576;
Best Local Similarity 61.0%; Pred. No. 9e-05;
Matches 25; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 EREKQMMREKELMLRLQDXEKTKKARELSEIQIRALQ 41
Db 333 EREKIEREKELMELKQIEQTKKAQOELEEQTRALE 373

RESULT 8
MOES_RAT
ID MOES_RAT STANDARD; PRT; 576 AA.
AC Q35763;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Moesin (Membrane-organizing extension spike protein).
GN MSN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
OX [1]
RN SEQUENCE FROM N.A., AND PHOSPHORYLATION.
RP MEDLINE=20404063; PubMed=10945828;
RA Theodorides T.C., Wang L., Pang X., Letourneau R., Culm K.E., Basu S.,
RA Wang Y., Correia I.;
RT "Cloning and cellular localization of the rat mast cell 78-kDa protein
RT phosphorylated in response to the mast cell 'stabilizer' cromolyn.";
RL J. Pharmacol. Exp. Ther. 294:810-821(2000).
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal
CC structures to the plasma membrane.

```

```

CC -!- PTM: Phosphorylated.
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF004811; AAB61666.1; -
CC InterPro; IPR000299; Band 4.1.
CC InterPro; IPR000798; Ez/rad/moesin.
CC Pfam; PF00373; Band 41; 1.
CC Pfam; PF00769; ERM_1.
CC PRINTS; PR00935; BAND41.
CC PROSITE; PS00660; FERM_1; 1.
CC PROSITE; PS00661; FERM_2; 1.
CC PROSITE; PS00507; FERM_3; 1.
CC Structural protein; Cytoskeleton; Phosphorylation.
CC KW Structural protein; Cytoskeleton.
CC FT INIT MET 0 BY SIMILARITY.
CC DOMAIN 1 294 FERM.
CC SEQUENCE 576 AA; 67607 MW; 59606907B2D89938 CRC64;

Query Match 65.3%; Score 126; DB 1; Length 576;
Best Local Similarity 61.0%; Pred. No. 9e-05;
Matches 25; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 EREKQMMREKELMLRLQDXEKTKKARELSEIQIRALQ 41
Db 333 EREKIEREKELMELKQIEQTKKAQOELEEQTRALE 373

RESULT 9
RADI_CHICK
ID RADI_CHICK STANDARD; PRT; 583 AA.
AC Q9PU45;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Radixin.
GN RDX.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=20225478; PubMed=10760599;
RA Li W., Crouch D.H.;
RT "Cloning and expression profile of chicken radixin.";
RL Biochim. Biophys. Acta 1491:327-332(2000).
CC -!- FUNCTION: Probably plays a crucial role in the binding of the
CC barbed end of actin filaments to the plasma membrane (By
CC similarity).
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ249838; CAB59977.1; -
CC InterPro; IPR000299; Band 4.1.
CC InterPro; IPR000798; Ez/rad/moesin.
CC Pfam; PF00373; Band 41; 1.
CC Pfam; PF00769; ERM_1.

```


[illegible]

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>) or send an email to license@isb-sib.ch.

EWBL; LO2320; AAA36541.1; -;
EWBL; BC047109; AAH47109.1; -;
PIR; A46127; A46127.
Aarhus/Ghent-2DPAGE; 2501; IEF.
Genew; HGNC:9944; RDX..
MIM; 179410; -;
InterPro; IPRO00299; Band 4.1.
InterPro; IPRO00798; EzRad/moesin.
InterPro; IPRO08954; Moesin.
Pfam; PF00373; Band_41; 1.
Pfam; PF00769; ERN; 1.
PRINTS; PR00935; BAND41.
SMART; SMC0295; B41; 1.
PROSITE; PS00660; FERM_1; 1.
PROSITE; PS00661; FERM_2; 1.
PROSITE; PS00037; FERM_3; 1.
Structural protein; Cytoskeleton; Actin-binding; Phosphorylation;
Actin capping.
DOMAIN 5 295 FERM.
DOMAIN 311 522 GLU-RICH.
DOMAIN 470 477 POLY-PRO.
CONFLICT 328 328 K->E (IN REF. 2).
SEQUENCE 583 AA; 68564 MW; 898697ELD675FFE7 CRC64;

Query Match 61.7%; Score 119; DE 1; Length 583;
Best Local Similarity 58.5%; Pred. No. 0.00039;
Matches 24; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Dy 1 EREKQWREKEELMLRLQDXEEKTKKAERELSFOIQRAIQ 41
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 334 EREKERIESEKELMERLKQIEEQITAKQELEEOTRKALE 374

RESULT 13
ID THRY_RABIT STANDARD; PRT; 1407 AA.
AC P37709;
CD 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
TM 10-OCT-2003 (Rel. 42, Last annotation update)
FT Trichohyalin.
DE
GN Oryctolagus cuniculus (Rabbit).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI TaxID=9986;
[1]
SQ SEQUENCE FROM NA.
NP Pietz M.J., Rogers G.E.;
AA Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
AL -! FUNCTION: Intermediate filament-associated protein that associates
CC in regular arrays with keratin intermediate filaments (KIF) of the
CC inner root sheath cells of the hair follicle and the granular
CC layer of the epidermis. It later becomes cross-linked to KIF by
CC isopeptide bonds. It may serve as scaffold protein, together
CC with involucrin, in the organization of the cell envelope or even
CC anchor the cell envelope to the KIF network. It may be involved in
CC its own calcium-dependent postsynthetic processing during terminal
CC differentiation.
CC -! SUBUNIT: Homodimer (Probable).
CC -! TISSUE SPECIFICITY: Found in the hard keratinizing tissues such as
CC the inner root sheath (IRS) of hair follicles and medulla, and in
CC the filiform papillae of dorsal tongue epithelium (Probable).
CC -! DEVELOPMENTAL STAGE: Expressed during late differentiation of
CC the epidermis.
CC -! DOMAIN: Consists of nine domains. Domain 1 contains two EF-hand
CC calcium-binding domains. Domains 2-4, 6, and 8 are almost entirely


```

FT REPEAT 707 742 1-11.
FT REPEAT 743 771 1-12.
FT REPEAT 772 796 1-13.
FT REPEAT 797 832 1-14.
FT DOMAIN 938 1507 23 X 23 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 938 961 2-1.
FT REPEAT 962 985 2-2.
FT REPEAT 986 1021 2-3.
FT REPEAT 1022 1044 2-4.
FT REPEAT 1045 1067 2-5.
FT REPEAT 1068 1090 2-6.
FT REPEAT 1091 1121 2-7.
FT REPEAT 1122 1144 2-8.
FT REPEAT 1145 1167 2-9.
FT REPEAT 1168 1197 2-10.
FT REPEAT 1198 1227 2-11.
FT REPEAT 1228 1250 2-12.
FT REPEAT 1251 1273 2-13.
FT REPEAT 1274 1296 2-14.
FT REPEAT 1297 1319 2-15.
FT REPEAT 1320 1342 2-16.
FT REPEAT 1343 1368 2-17.
FT REPEAT 1369 1391 2-18.
FT REPEAT 1392 1416 2-19.
FT REPEAT 1417 1439 2-20.
FT REPEAT 1440 1461 2-21.
FT REPEAT 1462 1484 2-22.
FT REPEAT 1485 1507 2-23.
FT VARSPLIC 1145 1197 Missing (in isoform Short).
FT VARSPLIC 1251 1273 /FTIDVSP 000847.
FT VARSPLIC 1399 1399 Missing (in isoform Short).
FT CONFLICT 1399 1399 E -> G (IN REF. 2).
SQ SEQUENCE 1549 AA; 201173 MW; E72PB9FF1326E54E CRC64;

Query Match 38.9%; Score 75; DB 1; Length 1549;
Best Local Similarity 40.5%; Pred. No. 9.8;
Matches 17; Conservative 10; Mismatches 9; Indels 6; Gaps 1;

QY 3 EKEQMMREKEELMLRLQD-----XEEKTKKAERLSEQIQR 38
Db 1348 EKEHLREREQQLRQEEGVFSQEEQLRRAEQEEQRRQR 1389

RESULT 15
TRHY_HUMAN STANDARD; PRT; 1898 AA.
AC Q07283;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Trichohyalin.
GN THH OR TRHY OR THL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93280194; PubMed=7685034;
RA Lee S.-C., Kim I.-G., Marekov L.N., O'Keefe E.J., Parry D.A.D.,
RA Steinert P.M.;
RT "The structure of human trichohyalin. Potential multiple roles as a
RT functional EF-hand-like calcium-binding protein, a cornified cell
RT envelope precursor, and an intermediate filament-associated (cross-
RT linking) protein.";
RL J. Biol. Chem. 268:12164-12176 (1993).
RN [2]
RP SEQUENCE OF 1731-1898 FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=93315897; PubMed=7686953;
RA O'Keefe E.J., Hamilton E.H., Lee S.-C., Steinert P.M.;
RT "Trichohyalin: a structural protein of hair, tongue, nail, and
RT epidermis.";

```

```

RL J. Invest. Dermatol. 101:65S-71S(1993).
CC -!- FUNCTION: Intermediate filament-associated protein that associates
CC in regular arrays with keratin intermediate filaments (KIF) of the
CC inner root sheath cells of the hair follicle and the granular
CC layer of the epidermis. It later becomes cross-linked to KIF by
CC isodipeptide bonds. It may serve as scaffold protein, together
CC with involucrin, in the organization of the cell envelope or even
CC anchor the cell envelope to the KIF network. It may be involved in
CC its own calcium-dependent postsynthetic processing during terminal
CC differentiation.
CC -!- SUBUNIT: Monomer (Probable).
CC -!- TISSUE SPECIFICITY: Found in the hard keratinizing tissues such as
CC the inner root sheath (IRS) of hair follicles and medulla, and in
CC the filiform papillae of dorsal tongue epithelium (Probable).
CC -!- DEVELOPMENTAL STAGE: Expressed during late differentiation of
CC the epidermis.
CC -!- DOMAIN: Consists of nine domains. Domain 1 contains two EF-hand
CC calcium-binding domains. Domains 2-4, 6, and 8 are almost
CC entirely alpha-helical, configured as a series of peptide repeats
CC of varying regularity, and are thought to form a single-stranded
CC alpha-helical rod stabilized by ionic interactions. Domain 6 is
CC the most regular and may bind KIF directly by ionic interactions.
CC Domains 5 and 7 are less well organized and may induce folds in
CC the molecule. Domain 9 contains the C-terminus, conserved among
CC different species.
CC -!- PTM: Substrate of transglutaminase. Some 200 arginines are
CC probably converted to citrullines by peptidylarginine deiminase.
CC -!- SIMILARITY: In the N-terminal section; belongs to the S-100
CC family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L09190; AAA65582.1; -.
CC PIR; A45973; A45973.
CC HSP; P02633; 4ICB.
CC Gene; HGNC:11791; THH.
CC MIM; 190370; -.
CC GO; GO:0005856; C:cytoskeleton; NAS.
CC GO; GO:0005509; F:calcium ion binding; TAS.
CC InterPro; IPR001751; CaBP_S100.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR002017; Spectrin.
CC Pfam; PF00036; ehand, 1.
CC Pfam; PF01023; S_100; 1.
CC ProDom; PD003407; CaBP_S100; 1.
CC PROSITE; PS00018; EF_HAND; 1.
CC PROSITE; PS00303; S100_CaBP; 1.
CC Keratinization; Calcium-binding; Repeat; Citrullination.
CC DOMAIN 1 91 S-100 LIKE.
CC CA_BIND 22 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CC CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CC DOMAIN 314 390 6 X 13 AA TANDEM REPEATS OF
CC REPEAT 314 326 R-R-E-Q-E-E-R-R-E-Q-Q-L.
CC REPEAT 327 339 1-1 (APPROXIMATE).
CC REPEAT 340 351 1-2 (APPROXIMATE).
CC REPEAT 352 364 1-3 (APPROXIMATE).
CC REPEAT 365 377 1-4.
CC REPEAT 378 390 1-5.
CC DOMAIN 391 444 9 X 6 AA TANDEM REPEATS OF R-R-E-Q-Q-L.
CC REPEAT 391 396 2-1.
CC REPEAT 397 402 2-2.
CC REPEAT 403 408 2-3.
CC REPEAT 409 414 2-4.
CC REPEAT 415 420 2-5.
CC REPEAT 421 426 2-6.

```

```
FT REPEAT 427 432 2-7.
FT REPEAT 433 438 2-8.
FT REPEAT 439 444 2-9.
FT DOMAIN 444 702 9 X 28 AA APPROXIMATE TANDEM REPEATS.
FT DOMAIN 923 1162 8 X 30 AA TANDEM REPEATS.
FT REPEAT 923 952 4-1.
FT REPEAT 953 982 4-2.
FT REPEAT 983 1012 4-3.
FT REPEAT 1013 1042 4-4.
FT REPEAT 1043 1072 4-5.
FT REPEAT 1073 1102 4-6.
FT REPEAT 1103 1132 4-7.
FT REPEAT 1133 1162 4-8.
FT DOMAIN 1250 1849 23 X 26 AA APPROXIMATE TANDEM REPEATS.
FT CONFLICT 1752 1752 F -> L (IN REF. 2).
FT CONFLICT 1794 1801 QERDQYR -> RSETGSTG (IN REF. 2).
FT CONFLICT 1857 1857 Q -> K (IN REF. 2).
FT CONFLICT 1880 1880 V -> G (IN REF. 2).
SQ SEQUENCE 1898 AA; 247219 MW; A74B5947FB62E31D CRC64;

Query Match 37.8%; Score 73; DB 1; Length 1898;
Best Local Similarity 41.2%; Pred. No. 18;
Matches 14; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 2 REKQMRKEELMLRLQDXEYKXARELSEQ 35
Db 1625 REEQQLQEGEQQLRQERDRKFRFEEQQLRQ 1658
|||:::| | | | | : | : | : |
|||:::| | | | | : | : | : |

Search completed: August 16, 2004, 09:44:10
Job time : 13 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:43:56 ; Search time 77 Seconds
(without alignments)
168.003 Million cell updates/sec

Title: US-09-856-070B-29

Perfect score: 193

Sequence: 1 EREKQMRKEELMLRLQD.....BEKTKARELSEIQRALQ 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	173	89.6	421	11	Q8CBU4 mus musculus
2	173	89.6	455	11	Q8VHK3 rattus norv
3	162	83.9	585	13	Q9YSW6 gallus gall
4	117	60.6	439	13	Q7ZWC7 brachydanio
5	110	57.0	580	13	Q9PT71 xenopus lae
6	104	53.9	159	4	Q9UJ26 homo sapien
7	102	52.8	158	4	Q9UJ27
8	92	47.7	156	4	Q9UJ28 homo sapien
9	92	47.7	161	4	Q9UK20
10	89	46.1	389	11	Q7TSG6
11	87	45.1	152	4	Q9UJ22
12	75	38.9	730	11	O88735
13	75	38.9	738	11	Q7TQL9
14	74.5	38.6	621	4	Q8N722 homo sapien
15	73.5	38.1	650	4	Q8N9M4
16	73	37.8	730	5	Q9VJ84 drosophila

17	73	37.8	788	11	Q8CF87	Q8cf87 mus musculu
18	73	37.8	851	11	Q8CF88	Q8cf88 mus musculu
19	71.5	37.0	456	4	Q8N824	Q8n824 homo sapien
20	71.5	37.0	503	5	Q25777	Q25777 plasmodium
21	71.5	37.0	585	5	Q960P3	Q960p3 drosophila
22	71.5	37.0	735	5	Q3UD5	Q9u0d5 tetrahymena
23	71.5	37.0	1168	5	Q9VT00	Q9vyu0 drosophila
24	71.5	37.0	3130	5	Q9BK46	Q9bk46 plasmodium
25	71.5	37.0	3130	5	Q8IDX6	Q8idx6 plasmodium
26	71	36.8	404	11	Q8CGD6	Q8cgd6 mus musculu
27	71	36.8	861	10	Q8H6S8	Q8hes8 pisum sativ
28	71	36.8	1048	11	Q8OTV1	Q8otv1 mus musculu
29	71	36.8	1100	4	Q9NV82	Q9nv82 homo sapien
30	71	36.8	1220	4	Q8NSA0	Q8nsa0 homo sapien
31	71	36.8	1298	4	Q7Z312	Q7z312 homo sapien
32	71	36.8	1299	4	Q7Z228	Q7z228 homo sapien
33	71	36.8	1329	4	Q75137	Q75137 homo sapien
34	70.5	36.5	1334	11	Q7TT13	Q7tt13 mus musculu
35	70.5	36.5	1418	5	Q86HT4	Q8ht4 dictyosteli
36	70	36.3	404	4	Q96CG5	Q96cg5 homo sapien
37	70	36.3	755	10	Q9FP71	Q9fp71 oryza sativ
38	70	36.3	1108	5	Q9ND10	Q9nd10 babesia big
39	70	36.3	2240	5	Q8I238	Q8ij38 plasmodium
40	69.5	36.0	581	11	Q7TSL7	Q7tsl7 mus musculu
41	69.5	36.0	1266	11	Q80TU6	Q80tu6 mus musculu
42	69.5	36.0	2962	5	Q93326	Q93326 caenorhabdi
43	69.5	36.0	3081	5	Q7YWF1	Q7ywf1 plasmodium
44	69	35.8	280	4	Q8NAK8	Q8nak8 homo sapien
45	69	35.8	365	10	Q8LDW7	Q8ldw7 arabidopsis

ALIGNMENTS

RESULT 1

Q8CBU4 PRELIMINARY; PRT; 421 AA.

ID Q8CBU4
AC Q8CBU4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Villin 2 (Fragment).
GN Vill2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMEL; AK035271; BAC29009.1; -.
DR MGD; MGI:98931; Vill2.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band_4.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR InterPro; IPR008954; Moesin.
DR Pfam; PF00373; Band_41; 1.
DR Pfam; PF00769; ERM; 1.
DR PROSITE; PS00661; PERM_2; 1.
DR PROSITE; PS50057; PERM_3; 1.
FT NON_TER
SQ SEQUENCE 421 AA; 50273 MW; E8D6C0D03F6E7497 CRC64;

Query Match 89.6%; Score 173; DB 11; Length 421;
Best Local Similarity 82.9%; Pred. No. 4.3e-10;
Matches 34; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

SO SEQUENCE 585 AA; 69366 MW; 854C6348F783CAEC CRC64;

Query Match	83.9%;	Score 162;	DB 13;	Length 585;
Best Local Similarity	78.0%;	Pred. No. 7.7e-09;		
Matches 32: Conservative	7;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY      1 BREKEQMREXFEELMLRLQDYEKTKKAERELSEQIORALQ 41  
        |||||:|||||:|:|:|:|:|:|:|:|:|:|:|:|  
        |||||:|||||:|:|:|:|:|:|:|:|:|:|:|:|  
DB     334 FREKEOMFREXFEELYLOEYEVKTAKRAELSDOIRALO 374
```

RESULT 4

PRELIMINARY;	PRT;	439 AA.
ID Q7ZWC7		
AC Q7ZWC7;		
8, 1993 (T-THRU)	24	00000001

DT 01-JUN-2003 (TRENBUREL. 24, Last sequence update)
DT 01-OCT-2003 (TRENBUREL. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.

CA NCBI_16616-1555,
RN [1]
RP SEQUENCE FROM N.A.

RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RD
RW EMBL: BC049479. A3449479 1. -
RZ

```

DR GO:0005556; C:cytoskeleton; IEA.
DR InterPro: IPR000299; Band 4.1.
DR InterPro: IPR000798; Ez/rad/moesin.

```

DR PFAM; PF00769; ERM; 1.
DR PRINTS; PR00935; BAND41.

```
DR PROSITE: PS00660; FERM_1; 1.
DR PROSITE: PS00661; FERM_2; 1.
DR PROSITE: PS50057; FERM_3; 1.
```

AM	hypothetical protein		
FT	NON TER	439	
SQ	SEQUENCE	439 AA; 52039 MW;	8C907EBB06D0A3D12 CRC64;

Query Match	60.6%;	Score 117;	DB 13;	Length 439;
Best Local Similarity	56.1%;	Pred. No. 0.00024;		
Matches	23;	Conservative	12;	Mismatches 6;
				Indels 0;
				Caps 0;

QY 1 EREXEQMREKEELMLRLQDYEEKTKKAERELSEQIQRALQ 41

[illegible]

Q9PT71	PRELIMINARY;	PRT;	580 AA.
ID	Q9PT71		
AC	Q9PT71		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Moedin.		
OS	Xenopus laevis	(African clawed frog).	
OC	Eukaryota;	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCBI_TaxID=8335;		
RN	[1]		
RP	SEQUENCE FROM N.A.		

RESULT 2
SQ SEQUENCE 585 AA; 69366 MW; 854C6348F783CAEC CRC64;

Q8VHK3	PRELIMINARY;	PRT;	455 AA.
ID	Q8VHK3		
AC	Q8VHK3;		
DT	01-MAR-2002	(M-EMBL)	20
Created			
Query Match	83.9%	Score	162;
Best Local Similarity	78.0%	Pred. No.	7.7e-09;
Matches	32;	Mismatches	2;
Conservative		Indels	0;
Gaps			0;

	Qy	1	BREKQMREKFEMLRLQDYEEKTKKAERLSBOIORALQ	41
DT	01-MAR-2002	(T=EMBLrel.	20,	Last sequence update)
DT	01-JUN-2003	(T=EMBLrel.	24,	Last annotation update)
DE	Ezrin	(Fragment).		
OR				
	Nb	334	BREKFMREKFEMLRVLOEVEVTKKAERLSBOIORALO	374

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI TaxID=10116;
RESULT 4

AN	[...]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Sprague-Dawley;

RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.

RL	EMBL; AF450298; AAL47844.1; --	IEA.
DR	GO; GO:0005856; C:cytoskeleton	
DR	InterPro; IPR000299; Band 4.1.	
DR	InterPro; IPR000798; Ez/rad/moesin.	
DR	Pfam; PF00373; Band 41.1.	
DR	Pfam; PF00059; ERM_1	
DR	Pfam; PF00059; ERM_1	

DR PRINTS; PR00935; EAND41.
DR SMART; SM00295; B41; 1.
DR PROSITE; PS00660; FERM 1; 1.
OK ACB1_16A01D-1955,
RN [1]
RP SEQUENCE FROM N.A.

DR	PROSITE;	PS00057;	FERM_3;	1.
FT	NON_TER	455	455	
RA	Strausberg R.;			
RL	Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.			
DE	EMBL: BC049479.1; EMBL: BC049479.1; EMBL: BC049479.1			

Query Match	Score 173;	DB 11;	Length 455;
Best Local Similarity	82.9%;	Pred. No. 4.6e-10;	
DR	GO:0005856; C:cytoskeleton; IEA.		
DR	GO:0005856; C:cytoskeleton; IEA.		
DR	InterPro; IPR000299; Band 4.1.		
DR	InterPro; IPR000798; Ez/rad/moesin.		

Matches	34; Conservative	6; Miscellaneous	1; Indexes	0; Maps	0; V, W, X, Y, Z, Misc	0; P, Q, R, S, T, U, V, W, X, Y, Z, Misc
Qy	1	ERKEQWMEKEBELMLRLQDSEKYNKAPRELSQIQRALQ	41			

334 EREKIQMLEKEBELMLRLQDFPQKTKRAEKELSENIEKALQ 374

	PRT;	585 AA.
ID Q9YGW6	PRELIMINARY;	SQ SEQUENCE 439 AA; 52039 MW; 8C907EB0D0A3D12 CRC64;
Q9YCW6		FT NON TER 439
RESULT 3		AN hypochromic protein.

DT	01-MAY-1999	(TREMblrel. 10, Created)	Query Match	60.6%	Score 117;	DB 13;	Length 439;
DT	01-MAY-1999	(TREMblrel. 10, Last sequence update)	Best Local Similarity	56.1%	Pred. No. 0.00024;		
DT	01-MAY-1999	(TREMblrel. 10, Last sequence update)	Mismatches	23;	Conservative	12;	Mismatches 6;
DT	01-MAY-1999	(TREMblrel. 10, Last sequence update)	Indels	0;			Gaps 0;

DE	Ezrin.
OS	Gallus gallus (Chicken).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	: ::: : ::: ::: ::: :::
	1EREKEQMWEEKELMLRLQDXEETKTKAERLSEIQIALQ41
	Qy

ACROSMAURIA; AVES; NEOGNATHIDAE; GALLIFORMES; FUSCIBALBES; FUSCIBALBES;
OC
Gallus.
OC
NCBI TaxID=9031;
OX

RP	SEQUENCE FROM N.A.	Q9PT71	PRELIMINARY;	PRT;	580 AA.
RC	STRAIN=White Leghorn; TISSUE=Brain;	AC	Q9PT71;		
RX	MEDLINE=99171943; PubMed=10511754;	DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
RA	Takahashi M., Yamagata M., Noda M.;	DT	01-MAR-2000 (TrEMBLrel. 13, Last sequence update)		
RT	"Specific expression of ezrin, a cytoskeletal-membrane linker protein,	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
RI	in a subset of chick retinotectal and sensory projections.";	DE	Moesin.		
RL	Eur. J. Neurosci. 11:545-558(1999).	OS	Xenopus laevis (African clawed frog).		
DR	ENBL; AB019790; BAA75497.1; -	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
DR	GO; GO:0005856; Cytoskeleton; IEA.	OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;		
DR	InterPro; IPR000289; Band_4.1.	OC	Xenopodinae; Xenopus.		
DR	InterPro; IPR000798; Ez/rad/moesin.	OX	NCBI_Taxid=8355;		
DR	InterPro; IPR008954; Moesin.	RN	[1] -		
DR	Pfam; PF00373; Band_4.1; 1.	RP	SEQUENCE FROM N.A.		
DR	Pfam; PF00769; ERN_1.				

RESULT 7	
Q9UJZ7	PRELIMINARY; PRT; 158 AA.
ID Q9UJZ7	
AC Q9UJZ7;	
DT 01-MAY-2000 (T-EMBLrel. 13, Created)	
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)	
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)	
DE Ezrin (Fragment).	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	

Qy	1	EREKEQMRFKEELMRLQ	19
Dd	138	EREKEQMRFKEELMRLQ	156
RESULT 9			
QSUKZ0		PRELIMINARY;	PRT; 161 AA.
AC	Q9UKZ0		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)	
DE	Ezzrin	(Fragment)	
OS	Homo sapiens	(Human).	
OC	Eukaryota;	Metazoa;	Chordata;
OC	Mammalia;	Eutheria;	Carnivora;
OX	NCBI TaxID=9606;		

```

RN RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA "Chen Z.C., Fadiel A., Naftolin F.;
RT "Mutation analysis of erin gene in cancer cells.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187552; AAD56713.1; -.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band 4.1.
DR Pfam; PF00373; Band_41; 1.
DR Pfam; PF00769; ERM_1; 1.
DR PROSITE; PS00557; FERM_3; 1.
FT NON TER 1
FT NON TER 161
SQ SEQUENCE 161 AA; 19439 MW; 5FD6EB910E017099 CRC64;

Query Match 47.7%; Score 92; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQMMREKELMLRLQ 19
DB 143 EREKQMMREKELMLRLQ 161

RESULT 10
Q7TSG6 PRELIMINARY; PRT; 389 AA.
AC Q7TSG6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053417; AHS3417.1; -.
KW Hypothetical protein.
SQ SEQUENCE 389 AA; 46367 MW; 6EA2EDCBA9BBSA76 CRC64;

Query Match 46.1%; Score 89; DB 11; Length 389;
Best Local Similarity 60.0%; Pred. No. 0.16;

```

```

Matches 18; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 EREKQMMREKELMLRLQDXBEKTKAER 30
DB 334 EKEKERIEREKELMERLQIEEQTVKAK 363

RESULT 11
Q9UJZ2 PRELIMINARY; PRT; 152 AA.
AC Q9UJZ2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome 2 (Fragment).
GN VIL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Chen Z.C., Fadiel A., Naftolin F.;
RT "Mutation of cytochrome 2 gene in glioblastoma.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF190059; AAF03158.1; -.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band_4.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR Pfam; PF00769; ERM; 1.
DR PROSITE; PS00557; FERM_3; 1.
FT NON TER 1
FT NON TER 152
FT NON TER 152
SQ SEQUENCE 152 AA; 18416 MW; B7FCE817DD2B9291 CRC64;

Query Match 45.1%; Score 87; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQMMREKELMLRL 18
DB 134 EREKQMMREKELMLRL 151

RESULT 12
O88735 PRELIMINARY; PRT; 730 AA.
AC O88735;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Microtubule-associated protein, MAP-115.
GN MAP7 OR MAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6CBA;
RX MEDLINE=98418230; PubMed=9745708;
RA Jonca J., Allaman J.M., Radlgruber G., Meda P., Kiss J.Z.,
RA French L.E., Masson D.;
RT "The distribution of murine 115-kDa epithelial microtubule-associated
RT protein (E-MAP-115) during embryogenesis and in adult organs suggests
RL a role in epithelial polarization and differentiation.";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Suzuki M.;
RT "Microtubule associated protein.";

```

```
RL EMBL; Y15197; CAA75495.1; --
DR EMBL; A3098611; BAC53729.1; --
DR MGD; MGI:1328328; Mtap7.
DR InterPro; IPR008604; E-MAP-115.
DR Pfam; PF05672; E-MAP-115; 1.
SQ SEQUENCE 730 AA; 82021 MW; F44B4201DB8F5A17 CRC64;

Query Match      38.9%; Score 75; DB 11; Length 730;
Best Local Similarity 48.6%; Pred. No. 7.7;
Matches 18; Conservative 8; Mismatches 9; Indels 2; Gaps 1;

QY 2 REKEQMMREKEELMLRLQDXEETKKAERLSEIQRLQ 38
Db 515 RLEEQAQREKEELALRL--AEERERWEREVERVQK 549

RESULT 13
Q7TQL9
ID Q7TQL9 PRELIMINARY; PRT; 738 AA.
AC Q7TQL9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to microtubule-associated protein 7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052637; AAH52637.1; --
SQ SEQUENCE 738 AA; 82839 MW; EFA2D1E1CC78815F CRC64;

Query Match      38.9%; Score 75; DB 11; Length 738;
Best Local Similarity 48.6%; Pred. No. 7.8;
Matches 18; Conservative 8; Mismatches 9; Indels 2; Gaps 1;

QY 2 REKEQMMREKEELMLRLQDXEETKKAERLSEIQRLQ 38
Db 523 RLEEQAQREKEELALRL--AEERERWEREVERVQK 557

RESULT 14
Q8N7Z2
ID Q8N7Z2 PRELIMINARY; PRT; 621 AA.
AC Q8N7Z2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ40198.
OS Homo sapiens (Human).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK097517; BAC05084.1; --
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
KW Hypothetical protein.
SQ SEQUENCE 621 AA; 77106 MW; 67938434F3BE7929 CRC64;

Query Match      38.6%; Score 74.5; DB 4; Length 621;
Best Local Similarity 35.3%; Pred. No. 7.5;
Matches 18; Conservative 11; Mismatches 11; Indels 11; Gaps 1;

QY 2 REKEQMMREKEELMLRLQDXEETKKAERLSEIQRLQ 41
Db 555 REQEMNQEQEERKWEQEKMCQEQEKMQEQEKMRQEQEKWQEVRLRQ 605

RESULT 15
Q8N9W4
ID Q8N9W4 PRELIMINARY; PRT; 650 AA.
AC Q8N9W4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ36144.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK093463; BAC04173.1; --
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
KW Hypothetical protein.
SQ SEQUENCE 650 AA; 79069 MW; 2DC6078E002D031 CRC64;

Query Match      38.1%; Score 73.5; DB 4; Length 650;
Best Local Similarity 42.5%; Pred. No. 9.9;
Matches 17; Conservative 9; Mismatches 11; Indels 3; Gaps 1;

QY 2 REKEQMMREKEELMLRLQDXEETKKAERLSEIQRLQ 41
Db 453 REREKKMRREBEETM--REQEKMKQKQEQENNWQEQEKWQ 489
```

us-09-856-070b-29.std.rspt

Mon Aug 16 10:46:54 2004

Search completed: August 16, 2004, 09:45:40
Job time : 78 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:50:52 ; Search time 55 Seconds

(without alignments)
210.626 Million cell updates/sec

Title: US-09-856-070b-29

Perfect score: 41

Sequence: 1 ERKEQVMREKELMLRLQD.....EEKTKAERLSEIQIRALQ 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 5

Total number of hits satisfying chosen parameters: 9586

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	48.8	34	4 AAB82020	Aab82020 Human hep
2	20	48.8	436	4 AAG73954	Aag73954 Human col
3	20	48.8	579	7 ADC31630	Adc31630 Human nov
4	20	48.8	585	7 ADE63987	Ade63987 Human pro
5	20	48.8	586	2 AAY27443	Aay27443 Amino aci
6	20	48.8	611	7 ADC31629	Adc31629 Human nov
7	20	48.8	622	4 AAU30004	Aau30004 Novel hum
8	20	48.8	628	7 ADC31631	Adc31631 Human nov
9	20	48.8	635	3 AAB53356	Aab53356 Human col
10	14	34.1	14	4 AAB82035	Aab82035 Human hep
11	12	29.3	52	4 AAU33060	Aau33060 Novel hum
12	10	24.4	13	4 AAB82037	Aab82037 Human hep
13	9	22.0	12	4 AAB82038	Aab82038 Human hep
14	8	19.5	8	4 AAU68025	Aau68025 Human bre
15	8	19.5	11	4 AAB82039	Aab82039 Human hep
16	8	19.5	349	3 AAB07823	Aab07823 Amino aci
17	7	17.1	32	4 AAB82019	Aab82019 Human hep
18	7	17.1	240	6 ABM69116	Abm69116 Photorhab
19	7	17.1	336	7 ADD22984	Add22984 Human pro
20	7	17.1	363	4 AAU00244	Aau00244 Human his
21	7	17.1	503	4 ABG16577	Abg16577 Novel hum
22	7	17.1	510	6 ABO15000	Abol15000 Human NOV
23	7	17.1	555	6 ABO15001	Abol15001 Human NOV
24	7	17.1	563	5 ABG61631	Abg61631 Streptoco
25	7	17.1	576	7 ADE54611	Ade54611 Human Pro

26	7	17.1	576	7	ADD45108	Add45108 Human pro
27	7	17.1	577	6	ABR59727	AbR59727 Human moe
28	7	17.1	577	7	ADB70362	AdB70362 Moesin SE
29	7	17.1	577	7	ADC77532	AdC77532 Human moe
30	7	17.1	583	6	ABU89709	Abu89709 Protein d
31	7	17.1	583	7	ADB79911	AdB79911 Mouse put
32	7	17.1	583	7	ADC56738	AdC56738 Murine ra
33	7	17.1	593	4	ABG19947	AbG19947 Novel hum
34	7	17.1	645	2	AAW22780	Aaw22780 human RAD
35	7	17.1	645	2	AAW71293	Aaw71293 Protein G
36	7	17.1	650	7	ADB65482	AdB65482 Human pro
37	7	17.1	667	3	AAU53008	Aau53008 Human sec
38	7	17.1	689	4	AAE01020	Aae01020 Human pif
39	7	17.1	742	3	AAB24238	Aab24238 Human pro
40	7	17.1	796	4	AAB93007	Aab93007 Human pro
41	7	17.1	817	4	AAB95640	Aab95640 Human pro
42	7	17.1	1215	5	AAE22856	Aae22856 Human pro
43	7	17.1	1215	5	AAU39661	Aau39661 Human his
44	7	17.1	1215	7	ADC21697	AdC21697 Human his
45	7	17.1	1272	7	ADD14124	Add14124 Human src

ALIGNMENTS

RESULT 1

AAB82020

ID AAB82020 standard; peptide; 34 AA.

AC AAB82020;

DT 13-JUN-2001 (first entry)

DE Human hepreceptor domain B.

KW Human; hepreceptor domain B; cytostatic; anti-HIV; antibiotic; nootropic;
KW immune response inducer; ezrin; infectious diseases; cancer;
KW HIV-related dementia.

OS Homo sapiens.

FX Key Location/Qualifiers

FT Modified-site 14

FT /note= "Optionally phosphorylated"

PN GB2354241-A.

PD 21-MAR-2001.

PF 17-SEP-1999; 99GB-00021881.

PR 17-SEP-1999; 99GB-00021881.

XX (HOLM/) HOLMS R D.

PI Holms RD;

DR WPI; 2001-293287/31.

XX Novel regulatory or unfolding peptides of ezrin that binds to

PT Hepreceptor, useful for inducing immune response for treating infectious

PT diseases and cancer.

XX Claim 5; Page 36; 42pp; English.

PS The present sequence is domain B of human hepreceptor of human ezrin. The
XX hepreceptor is a novel active site in human ezrin. Ezrin regulates the
CC structure of the cortical cytoskeleton to control cell surface
CC topography. The present invention relates to peptides (see AAB82021 to
CC AAB82041) that bind to hepreceptor with greater affinity than HEPI (see
CC AAB82046). The hepreceptor binding peptides are useful for inducing
CC immune response, and for treating infectious diseases, cancer and HIV-
CC related dementia. The present sequence assemblies into two anti-parallel

CC helices with hepreceptor domain A (see AAB82019)

XX Sequence 34 AA;

Query Match 48.8%; Score 20; DB 4; Length 34;

Best Local Similarity 100.0%; Pred. No. 1.9e-12;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 EEKTKAARELSGQIQALQ 41

DB 15 EEKTKAARELSGQIQALQ 34

RESULT 2

AAG73954

ID AAG73954 standard; protein; 436 AA.

XX AC

XX AAG73954;

XX DT

03-SEP-2001 (first entry)

XX DE

Human colon cancer antigen protein SEQ ID NO:4718.

XX KW

Human; colon cancer; colon cancer antigen; diagnosis; detection; colorectal carcinoma.

XX OS

Homo sapiens.

XX PN

WO200122920-A2.

XX PD

05-APR-2001.

XX PF

28-SEP-2000; 2000WO-US026524.

XX PR

29-SEP-1999; 99US-0157137P.

XX PR

03-NOV-1999; 99US-0163280P.

XX PA

(HUMA-) HUMAN GENOME SCI INC.

XX PI

Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI

2001-235357/24.

XX DR

N-PSDB; AAH33385.

XX PT

Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers.

XX PS

Claim 11; Page 6520-6521; 9803pp; English.

XX CC

AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patient's own production of P. Additionally, N may be used to produce the colon cancer-associated Ps, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent sequences used in the exemplification of the present invention. N.B. Pages 666 to 692 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922

XX SQ

Sequence 436 AA;

Query Match

Best Local Similarity 100.0%; Pred. No. 1.7e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKEQWMREKEELMLRLQD 20

DB 184 EREKEQWMREKEELMLRLQD 203

RESULT 3

ADC31630

ID ADC31630 standard; protein; 579 AA.

XX AC

ADC31630;

XX DT

18-DEC-2003 (first entry)

XX DE

Human novel polypeptide sequence, SEQ ID NO:1712.

XX KW

Human; diagnostic; drug screening; forensics; gene mapping; biodiversity assessment; Parkinson's disease; Alzheimer's disease; neurodegenerative diseases; anaemia; platelet disorder; wound; burns; ulcers; osteoporosis; autoimmune disease; cancer; molecular weight marker; food supplement; antiparkinsonian; nootropic; neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary; antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic; gene therapy; chromosome 6q25.2-26.

XX OS

Homo sapiens.

XX PN

WO2003029271-A2.

XX PD

10-APR-2003.

XX PF

24-SEP-2002; 2002WO-US030474.

XX PR

24-SEP-2001; 2001US-0324631P.

XX PA

(HYSE-) HYSEQ INC.

XX PI

Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

XX PI

Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

XX PI

Haley-Vicente D, Drmanac RT;

XX WPI

2003-371981/35.

XX DR

N-PSDB; ADC30659.

XX PT

New polynucleotide and polypeptide useful for diagnosing, preventing or treating conditions such as neurodegenerative diseases, anemias, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer.

XX PS

Claim 20; SEQ ID NO 1712; 1185pp; English.

XX CC

The invention relates to 971 novel human cDNA sequences (ADC99919-ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The invention also relates to nucleic acid sequences over 99% identical with the novel human cDNAs. The invention additionally encompasses expression vectors and host cells comprising a nucleic acid of the invention; the recombinant production of a polypeptide of the invention; a method of detecting against a polypeptide of the invention; a method of detecting polynucleotides or polypeptides of the invention; and methods of identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition; kits comprising polynucleotide probes and/or monoclonal antibodies for carrying out the methods of the invention; methods for the identification of compounds that modulate the expression or activity of the polynucleotide and/or polypeptide; and 767 cDNA sequences corresponding to the cDNA sequences of the invention (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628-ADC33394). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's

CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 579 AA;

Query Match 48.8%; Score 20; DB 7; Length 579;

Best Local Similarity 100.0%; Pred. No. 2.2e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EREKEQMMREKEELMLRLQD 20

Db 302 EREKEQMMREKEELMLRLQD 321

RESULT 4

AD663987

ID AD663987 standard; protein; 585 AA.

XX AC AD663987;

XX 29-JAN-2004 (first entry)

XX Human Protein P15311, SEQ ID NO 9933.

XX Human; pain; neuronal tissue; gene therapy;

XX spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P15311.

XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 585 AA;

Query Match 48.8%; Score 20; DB 7; Length 585;

Best Local Similarity 100.0%; Pred. No. 2.3e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EREKEQMMREKEELMLRLQD 20

Db 333 EREKEQMMREKEELMLRLQD 352

RESULT 5

RAY27443

ID RAY27443 standard; protein; 586 AA.

XX AC RAY27443;

XX 26-NOV-1999 (first entry)

XX Amino acid sequence of human ezrin polypeptide.

XX Pharmaceutical; ezrin; mutant; tumor; metastasis; human.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 354

XX /notes "the Tyr at this position can be mutated
XX (preferably to a Phe) to construct an ezrin mutant of the
XX invention"

XX WO9947150-A2.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-EF002054.

XX 18-MAR-1998; 98US-00040725.

XX (CURI-) INST CURIE.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Arpin M, Crepaldi T, Gautreau A, Louvard D;

XX WPI; 1999-561851/47.

XX New composition for prevention and treatment of tumors and metastasis.

XX Example 1; Fig 1; 31pp; English.

XX The invention provides a pharmaceutical composition containing ezrin
XX protein, RNA or DNA mutated on tyrosine 353, or a functional fragment or
XX derivative of the ezrin mutant. The new composition is useful for
XX prevention and/or treatment of tumors, and especially metastasis. The
XX present sequence represents the amino acid sequence of human ezrin
XX (before the maturation by deletion of the first amino acid Met)

SQ Sequence 586 AA;

Query Match 48.8%; Score 20; DB 2; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.3e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKEQWMREKELMLRLQD 20
|||
Db 334 EREKEQWMREKELMLRLQD 353

RESULT 6

AD31629

ID ADC31629 standard; protein; 611 AA.

XX

AC ADC31629;

XX

DT 18-DEC-2003 (first entry)

XX Human novel polypeptide sequence, SEQ ID NO:1711.

DE

XX Human; diagnostic; drug screening; forensics; gene mapping;

XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cyrostatic;

KW gene therapy; chromosome 6q25.2-26.

XX

OS Homo sapiens.

XX

PN WO2003029271-A2.

XX

XX 10-APR-2003.

XX

XX 24-SEP-2002; 2002WO-US030474.

XX

XX 24-SEP-2001; 2001US-0324631P.

PR

XX (HYSE-) HYSEQ INC.

PA

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

PI Haley-Vicente D, Drmanac RT;

XX

XX WPI; 2003-371981/35.

DR N-PSDB; ADC30658.

XX

XX New polynucleotide and polypeptide useful for diagnosing, preventing or

PT treating conditions such as neurodegenerative diseases, anemias, platelet

PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

PT cancer.

XX

XX Claim 20; SEQ ID NO 1711; 1185pp; English.

PS

XX The invention relates to 971 novel human cDNA sequences (ADC29919-

CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The

CC invention also relates to nucleic acid sequences over 99% identical with

CC the novel human cDNAs. The invention additionally encompasses expression

CC vectors and host cells comprising a nucleic acid of the invention; the

CC recombinant production of a polypeptide of the invention; an antibody

CC against a polypeptide of the invention; a method of detecting

CC polynucleotides or polypeptides of the invention; and methods of

CC identifying a compound which binds to a polypeptide of the invention. The

CC invention further discloses methods of preventing, treating or

CC ameliorating a medical condition; kits comprising polynucleotide probes

CC and/or monoclonal antibodies for carrying out the methods of the

CC invention; methods for the identification of compounds that modulate the

CC expression or activity of the polynucleotide and/or polypeptide; and 767

CC contig sequences corresponding to the cDNA sequences of the invention

CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628

CC -ADC33394). The nucleic acids and polypeptides of the invention are

CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anemias, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 611 AA;

Query Match 48.8%; Score 20; DB 7; Length 611;

Best Local Similarity 100.0%; Pred. No. 2.3e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EREKEQWMREKELMLRLQD 20
|||
Db 334 EREKEQWMREKELMLRLQD 353

RESULT 7

AAU30004

ID AAU30004 standard; protein; 622 AA.

XX

AC AAU30004;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secreted protein #495.

XX

XX Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX

OS Homo sapiens.

XX

PN WO200179449-A2.

XX

XX 25-OCT-2001.

XX

XX 16-APR-2001; 2001WO-US008656.

XX

XX 18-APR-2000; 2000US-00552929.

PR

XX 26-JAN-2001; 2001US-00770160.

XX

XX (HYSE-) HYSEQ INC.

PA

XX Tang YT, Liu C, Drmanac RT;

PI

XX WPI; 2001-611725/70.

XX

XX Nucleic acids encoding a range of human polypeptides, useful in genetic

CC vaccination, testing and therapy.

CC

XX Claim 20; Page 219; 765pp; English.

XX

XX The invention relates to novel human secreted polypeptides. The

CC polypeptides and antibodies to the polypeptides are useful for

CC determining the presence of or predisposition to a disease associated

CC with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells

CC expressing the proteins (agonists and antagonists) are useful for

CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising the

CC nucleic acids encoding the polypeptides and cells genetically engineered

CC to express them are also useful for producing the proteins. The proteins

CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation, to regulate haematopoiesis, and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
XX Sequence 622 AA;
SQ

Query Match 48.8%; Score 20; DB 4; Length 622;
Best Local Similarity 100.0%; Pred. No. 2.4e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQWREKEELMLRLQD 20
| | | | | | | | | | | | | | | | | | | | | |
DB 370 EREKQWREKEELMLRLQD 389

RESULT 8
ADC31631
ID ADC31631 standard; protein; 628 AA.
XX
XX ADC31631;
AC
XX
XX 18-DEC-2003 (first entry)
XX
XX Human novel polypeptide sequence, SEQ ID NO:1713.
DE
XX Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnery;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 6q25.2-26.
XX
XX Homo sapiens.
OS
XX WO2003029271-A2.
PN
XX 10-APR-2003.
PD
XX 24-SEP-2002; 2002WO-US030474.
PF
XX 24-SEP-2001; 2001US-0324631P.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
PI
XX WPI; 2003-371981/35.
DR
XX N-PSDB; ADC30660.
DR
XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
PT
PT
PS Claim 20; SEQ ID NO 1713; 1185pp; English.
PS
XX

CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The

CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC3394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 628 AA;
SQ

Query Match 48.8%; Score 20; DB 7; Length 628;
Best Local Similarity 100.0%; Pred. No. 2.4e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQWREKEELMLRLQD 20
| | | | | | | | | | | | | | | | | | | | | |
DB 351 EREKQWREKEELMLRLQD 370

RESULT 9
AAB53356
ID AAB53356 standard; protein; 635 AA.
XX
XX AAB53356;
AC
XX
XX 09-MAR-2001 (first entry)
DT
XX Human colon cancer antigen protein sequence SEQ ID NO:896.
DE
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnery;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; Gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder.
XX
XX Homo sapiens.
OS
XX WO200055351-A1.
PN
XX 21-SEP-2000.
PD
XX 08-MAR-2000; 2000WO-US005883.
PF
XX 12-MAR-1999; 99US-0124270P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI; 2000-587534/55.
DR
XX N-PSDB; AAC98113.
DR
XX Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer.

QY 5 EQWMEKEELML 16
Db 6 EQWMEKEELML 17

RESULT 12

AAB82037
ID AAB82037 standard; peptide; 13 AA.

XX AC AAB82037;

XX DT 13-JUN-2001 (first entry)

XX DE Human hepreceptor domain A binding peptide Rupe2032.

XX KW Human; hepreceptor; cytostatic; anti-HIV; antibiotic; nootropic;
KW immune response inducer; ezrin; infectious diseases; cancer;
KW HIV-related dementia.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 11
FT /note= "Optionally phosphorylated"

XX GN GB2354241-A.

XX PD 21-MAR-2001.

XX PF 17-SEP-1999; 99GB-00021881.

XX PR 17-SEP-1999; 99GB-00021881.

XX PA (HOLM/) HOLMS R D.

XX PI Holms RD;

XX DR WPI; 2001-293287/31.

XX PT Novel regulatory or unfolding peptides of ezrin that binds to
PT hepreceptor, useful for inducing immune response for treating infectious
PT diseases and cancer.

XX PS Claim 22; Page 36; 42pp; English.

XX CC The hepreceptor is a novel active site in human ezrin. Ezrin regulates
CC the structure of the cortical cytoskeleton to control cell surface
CC topography. The present invention relates to peptides (see AAB82021 to
CC AAB82041) that bind to hepreceptor with greater affinity than HEP1 (see
CC AAB82046). The hepreceptor binding peptides are useful for inducing
CC immune response, and for treating infectious diseases, cancer and HIV-
CC related dementia. The present peptide binds to domain A of the
CC hepreceptor (AAB82019)

XX SQ Sequence 13 AA;

Query Match 24.4%; Score 10; DB 4; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 KEELMLRLQD 20

Db 1 KEELMLRLQD 10

RESULT 13

AAB82038
ID AAB82038 standard; peptide; 12 AA.

XX AC AAB82038;

XX DT 13-JUN-2001 (first entry)

XX XX

DE Human hepreceptor domain A binding peptide Rupe2132.

XX KW Human; hepreceptor; cytostatic; anti-HIV; antibiotic; nootropic;
KW immune response inducer; ezrin; infectious diseases; cancer;
KW HIV-related dementia.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 10
FT /note= "Optionally phosphorylated"

XX GN GB2354241-A.

XX PD 21-MAR-2001.

XX PF 17-SEP-1999; 99GB-00021881.

XX PR 17-SEP-1999; 99GB-00021881.

XX PA (HOLM/) HOLMS R D.

XX PI Holms RD;

XX DR WPI; 2001-293287/31.

XX PT Novel regulatory or unfolding peptides of ezrin that binds to
PT hepreceptor, useful for inducing immune response for treating infectious
PT diseases and cancer.

XX PS Claim 24; Page 36; 42pp; English.

XX CC The hepreceptor is a novel active site in human ezrin. Ezrin regulates
CC the structure of the cortical cytoskeleton to control cell surface
CC topography. The present invention relates to peptides (see AAB82021 to
CC AAB82041) that bind to hepreceptor with greater affinity than HEP1 (see
CC AAB82046). The hepreceptor binding peptides are useful for inducing
CC immune response, and for treating infectious diseases, cancer and HIV-
CC related dementia. The present peptide binds to domain A of the
CC hepreceptor (AAB82019)

XX SQ Sequence 12 AA;

Query Match 22.0%; Score 9; DB 4; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 EELMLRLQD 20

Db 1 EELMLRLQD 9

RESULT 14

AAU68025
ID AAU68025 standard; peptide; 8 AA.

XX AC AAU68025;

XX DT 16-JAN-2002 (first entry)

XX DE Human Breast cancer-associated protein isoform, BPI-1 peptide #3.

XX KW Human; Breast cancer-associated protein isoform; breast cancer;
KW immunogen; cytostatic; BPI; tryptic digest peptide.

XX OS Homo sapiens.

XX GN WO200171357-A2.

XX PD 27-SEP-2001.

XX PF 20-MAR-2001; 2001WO-GB001219.

XX XX

PR 20-MAR-2000; 2000GB-00006695.
PR 24-MAR-2000; 2000GB-00007265.

PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

PI Herath HMAC, O'hare MJ, Page MJ, Parekh RB, Waterfield MD;

XX WPI; 2001-611532/70.

XX Identifying proteins for clinical screening, diagnosis and prognosis of
PT breast cancer, comprises detecting Breast Cancer-Associated Protein
PT isoforms (BPIs) using two-dimensional electrophoresis.

XX Claim 9; Page 43; 197pp; English.

XX The invention relates to diagnosing, determining the stage or severity,
CC or identifying the risk of a subject developing cancer (especially breast
CC cancer), or monitoring the effect of therapy on a subject with cancer
CC comprising analysing a test sample using two-dimensional electrophoresis
CC and detecting Breast Cancer-Associated Protein Isoforms (BPIs). The
CC methods disclosed are used for the diagnosis and prognosis of breast
CC cancer, for determining the severity of breast cancer, and for
CC identifying a subject at risk of developing breast cancer, and monitoring
CC the effect of therapy administered to a subject. Antibodies raised
CC against the binding domain of a BPI, the binding domain of a BPI, a
CC nucleic acid encoding a BPI, or a nucleic acid that inhibits the function
CC of a BPI can be incorporated into a pharmaceutical composition for
CC treating or preventing breast cancer. The methods use sensitive and
CC specific biomarkers provide early diagnosis of breast cancer, and the
CC compositions are more potent, specific, and has a more rapid effect with
CC fewer side effects than other prior art methods. The present sequence is
CC a tryptic digest peptide from a BPI of the invention

XX SQ Sequence 8 AA;

Query Match 19.5%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 EKELMLR 17
Db 1 EKELMLR 8

RESULT 15

AAB82039
ID AAB82039 standard; peptide; 11 AA.

XX AC AAB82039;

DT 13-JUN-2001 (first entry)

DE Human hepreceptor domain A binding peptide Rupe2232.

XX Human; hepreceptor; cytostatic; anti-HIV; antibiotic; nootropic;
KW immune response inducer; ezrin; infectious diseases; cancer;
KW HIV-related dementia.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 9 /note= "Optionally phosphorylated"

XX GB2354241-A.

XX PD 21-MAR-2001.

XX PF 17-SEP-1999; 99GB-00021881.

XX PR 17-SEP-1999; 99GB-00021881.

XX (HOLM/) HOLMS R D.

XX Holms RD;

XX WPI; 2001-293287/31.

XX Novel regulatory or unfolding peptides of ezrin that binds to
PT hepreceptor, useful for inducing immune response for treating infectious
PT diseases and cancer.

XX Claim 26; Page 37; 42pp; English.

XX The hepreceptor is a novel active site in human ezrin. Ezrin regulates
CC the structure of the cortical cytoskeleton to control cell surface
CC topography. The present invention relates to peptides (see AAB82021 to
CC AAB82041) that bind to hepreceptor with greater affinity than HEP1 (see
CC AAB82046). The hepreceptor binding peptides are useful for inducing
CC immune response, and for treating infectious diseases, cancer and HIV-
CC related dementia. The present peptide binds to domain A of the
CC hepreceptor (AAB82019)

XX SQ Sequence 11 AA;

Query Match 19.5%; Score 8; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ELMRLQD 20
Db 1 ELMRLQD 8

Search completed: August 16, 2004, 09:54:33

Job time : 56 secs

GenCore version 5.1.6.
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:54:23 ; Search time 39 Seconds
(without alignments)

54.273 Million cell updates/sec

Title: US-09-856-070B-29

Perfect score: 41

Sequence: 1 EREKEQMMREKEELMLRLQD.....EEKTKAERISEQIQALQ 41

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 5

Total number of hits satisfying chosen parameters: 1627

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	48.8	586	4	US-09-040-725A-1
2	10	24.4	27	4	US-09-040-725A-2
3	7	17.1	249	4	US-09-543-681A-8153
4	7	17.1	272	4	US-09-489-039A-14059
5	7	17.1	336	4	US-09-848-294-5
6	7	17.1	645	2	US-08-592-126-144
7	7	17.1	645	2	US-08-687-080-47
8	7	17.1	645	4	US-09-168-595-144
9	7	17.1	689	4	US-09-425-335-2
10	7	17.1	1312	2	US-08-592-126-148
11	7	17.1	1312	2	US-08-687-080-51
12	7	17.1	1312	4	US-09-168-595-148
13	6	14.6	12	2	US-08-538-711A-6
14	6	14.6	12	3	US-08-725-027-6
15	6	14.6	12	4	US-09-542-552-6
16	6	14.6	17	4	US-09-589-462-1
17	6	14.6	29	1	US-08-324-301-6
18	6	14.6	101	4	US-09-821-976-4910
19	6	14.6	160	4	US-09-489-039A-11628
20	6	14.6	174	4	US-09-500-569-26
21	6	14.6	174	4	US-09-971-823B-26
22	6	14.6	196	4	US-09-711-164-379
23	6	14.6	276	4	US-09-328-352-7013
24	6	14.6	282	1	US-08-324-301-15
25	6	14.6	287	3	US-09-105-697-7
26	6	14.6	291	3	US-09-105-697-5
27	6	14.6	305	4	US-09-134-000C-6304

28	6	14.6	316	4	US-09-134-001C-3736	Sequence 3736, Ap
29	6	14.6	327	3	US-08-960-780-23	Sequence 23, Appl
30	6	14.6	327	3	US-08-960-780-29	Sequence 29, Appl
31	6	14.6	327	3	US-09-073-898-23	Sequence 23, Appl
32	6	14.6	327	3	US-09-073-898-29	Sequence 29, Appl
33	6	14.6	327	4	US-09-850-351A-23	Sequence 23, Appl
34	6	14.6	327	4	US-09-850-351A-29	Sequence 29, Appl
35	6	14.6	341	2	US-08-538-711A-8	Sequence 8, Appl
36	6	14.6	341	3	US-08-725-027-8	Sequence 8, Appl
37	6	14.6	341	4	US-09-542-552-8	Sequence 14, Appl
38	6	14.6	362	4	US-09-500-569-14	Sequence 14, Appl
39	6	14.6	362	4	US-09-971-823B-14	Sequence 3, Appl
40	6	14.6	383	2	US-09-031-392-3	Sequence 3, Appl
41	6	14.6	383	3	US-09-299-549-3	Sequence 3, Appl
42	6	14.6	421	4	US-09-610-417-3	Sequence 4030, Ap
43	6	14.6	421	4	US-09-107-532A-4030	Sequence 13, Appl
44	6	14.6	434	1	US-08-111-939-13	Sequence 7, Appl
45	6	14.6	434	3	US-09-233-989-7	

ALIGNMENTS

RESULT 1
US-09-040-725A-1
; Sequence 1, Application US/09040725A
; Patent No. 6399584
; GENERAL INFORMATION:
; APPLICANT: Institut Curie
; APPLICANT: CNRS
; APPLICANT: Arpin, Monique
; APPLICANT: Crepaldi, Tiziana
; APPLICANT: Gautreau, Alexis
; APPLICANT: Louvard, Daniel
; TITLE OF INVENTION: Pharmaceutical composition containing ezrin mutated
; FILE REFERENCE: 39108200100
; CURRENT APPLICATION NUMBER: US/09/040,725A
; CURRENT FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-040-725A-1

Query Match 48.8%; Score 20; DB 4; Length 586;
Best Local Similarity 100.0%; Pred. NO. 2.6e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKEQMMREKEELMLRLQD 20
| | | | | | | | | | | | | | | | | | | | | |
DB 334 EREKEQMMREKEELMLRLQD 353

RESULT 2
US-09-040-725A-2
; Sequence 2, Application US/09040725A
; Patent No. 6399584
; GENERAL INFORMATION:
; APPLICANT: Institut Curie
; APPLICANT: CNRS
; APPLICANT: Arpin, Monique
; APPLICANT: Crepaldi, Tiziana
; APPLICANT: Gautreau, Alexis
; APPLICANT: Louvard, Daniel
; TITLE OF INVENTION: Pharmaceutical composition containing ezrin mutated
; FILE REFERENCE: 39108200100
; CURRENT APPLICATION NUMBER: US/09/040,725A
; CURRENT FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 4

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 27
; TYPE: PRN
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (22)
; OTHER INFORMATION: Xaa = ty
US-09-040-725A-2

```

Query Match 24.4%; Score 10; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 11; Conservative 0; Mismatches 0; Indels

Qy 16 LRLQDXEETK 26
17 LRLQDXEETK 27

RESULT 3
US-09-543-681A-8153
; Sequence 8153, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:

```
Query Match      17.1%; Score 7; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels
```

Qy 28 AERELSE 34
db 43 AERELSE 49

```

RESULT 4
US-09-489-039A-14059
/ Sequence 14059, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOS
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 14059
/ LENGTH: 272
/ TYPE: PRT
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14059

```

Query Match 17.1%; Score 7; DB 4; Length 272;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels

Qy	28	AER	34
D _b	66	AER	72

RESULT 5
US-09-848-294-5
; Sequence 5, Application US/09848294

```

1  APPLICANT: Tomkars, Nicholas K.
2  TITLE OF INVENTION: Isolation of A cDNA Encoding A No. 6479640el
3  TITLE OF INVENTION: Protein Tyrosine Phosphatase Which Localizes to Focal
4  TITLE OF INVENTION: Adhesions and Uses Therefor
5  FILE REFERENCE: CSHL90-04FZA
6  CURRENT APPLICATION NUMBER: US/09/848,294
7  CURRENT FILING DATE: 2001-05-03
8  PRIOR APPLICATION NUMBER: 09/235,251
9  PRIOR FILING DATE: 1999-01-22
10 PRIOR APPLICATION NUMBER: 08/759,536
11 PRIOR FILING DATE: 1996-12-04
12 PRIOR APPLICATION NUMBER: 08/107,420
13 PRIOR FILING DATE: 1993-08-16
14 PRIOR APPLICATION NUMBER: 07/663,579
15 PRIOR FILING DATE: 1991-03-01
16 PRIOR APPLICATION NUMBER: 07/494,036
17 PRIOR FILING DATE: 1990-03-14
18 NUMBER OF SEQ ID NOS: 13
19 SOFTWARE: FastSeq for Windows Version 3.0
20 SEQ ID NO 5

```

Query Match	17.1%;	Score 7;	DB 4;	Length 336;
Best Local Similarity	100.0%;	Pred. No. 14;		
Matches 7: Conservative	0;	Mismatches	0;	Indels

Qy 1 EREKEQM 7
Db 330 EREKEQM 336

```

RESULT 6
US-08-592-126-144
; Sequence 144, Application US/08592126
; Patent No. 5821091
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,126
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 144:
SEQUENCE CHARACTERISTICS:
LENGTH: 645 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G18.pep
US-08-592-126-144

Query Match 17.1%; Score 7; DB 2; Length 645;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KAERELS 33
|||||
DB 483 KAERELS 489

RESULT 7
US-08-687-080-47
Sequence 47, Application US/08687080
Patent No. 5965427
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSER: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,080
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/592,126
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 645 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G18.pep
US-08-687-080-47

Query Match 17.1%; Score 7; DB 2; Length 645;

Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KAERELS 33
|||||
DB 483 KAERELS 489

RESULT 8
US-09-168-595-144
Sequence 144, Application US/09168595
Patent No. 655666
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
Polyptides
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSER: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/168,595
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,126
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 144:
SEQUENCE CHARACTERISTICS:
LENGTH: 645 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G18.pep
US-09-168-595-144

Query Match 17.1%; Score 7; DB 4; Length 645;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KAERELS 33
|||||
DB 483 KAERELS 489

RESULT 9
US-09-425-335-2
Sequence 2, Application US/09425335
Patent No. 6518052
GENERAL INFORMATION:
APPLICANT: WEINMANN, ROBERTO
TITLE OF INVENTION: HUMAN HOMOLOGUE OF YEAST HELICASE AND USES THEREOF
FILE REFERENCE: db7 sequence

```

; CURRENT APPLICATION NUMBER: US/09/425,335
; CURRENT FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 689
; TYPE: PRT
; ORGANISM: HUMAN
US-09-425-335-2

Query Match      17.1%; Score 7; DB 4; Length 689;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ELMRLQ 19
Db 52 ELMRLQ 58

RESULT 10
US-08-592-126-148
; Sequence 148, Application US/08592126
; Patent No. 5821091
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,126
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: TRANS. OF RAD50 CDNA (SEQ. 54), NT.
; INDIVIDUAL ISOLATE: 389 TO 4324
; US-08-687-080-51

Query Match      17.1%; Score 7; DB 2; Length 1312;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KAERELS 33
Db 483 KAERELS 489

RESULT 12
US-09-168-595-148
; Sequence 148, Application US/09168595
; Patent No. 6555666
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; US-08-687-080-51

Query Match      17.1%; Score 7; DB 2; Length 1312;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KAERELS 33
Db 483 KAERELS 489

RESULT 11
US-08-687-080-51
```



```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/168,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,126
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
; US-09-168-595-148

```

```

Query Match 17.1%; Score 7; DB 4; Length 1312;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 27 KAERELS 33
DB 483 KAERELS 489

```

```

RESULT 13
US-08-538-711A-6
; Sequence 6, Application US/08538711A
; Patent No. 5994062
; GENERAL INFORMATION:
; APPLICANT: MULSHINE, JAMES, L.
; TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND
; TITLE OF INVENTION: DNA THEREOF FOR USE IN EARLY CANCER DETECTION
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,711A
; FILING DATE: 02-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34,556
; REFERENCE/DOCKET NUMBER: 2026-4201
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 6:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 12
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: peptide
; US-08-538-711A-6
Query Match 14.6%; Score 6; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EREKEQ 6
DB 1 EREKEQ 6

```

```

RESULT 14
US-08-725-027-6
; Sequence 6, Application US/08725027
; Patent No. 6251586
; GENERAL INFORMATION:
; APPLICANT: MULSHINE, JAMES, L.
; APPLICANT: TUCKMAN, MELVYN, S.
; TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND
; TITLE OF INVENTION: DNA THEREOF FOR USE IN EARLY CANCER DETECTION
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725,027
; FILING DATE: 02-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US08/538,711
; FILING DATE: 02-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34,556
; REFERENCE/DOCKET NUMBER: 2026-4201US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: peptide
; US-08-725-027-6

```

```

Query Match 14.6%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EREKEQ 6
DB 1 EREKEQ 6

```

```

RESULT 15
US-09-542-552-6
; Sequence 6, Application US/09542552

```

Patent No. 6500625
GENERAL INFORMATION:
APPLICANT: MULSHINE, JAMES. L.
TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND
METHOD OF INVENTION: DNA THEREOF FOR USE IN EARLY CANCER DETECTION
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/542,552
FILING DATE: 03-APR-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/538,711
FILING DATE: 02-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34,556
REFERENCE/DOCKET NUMBER: 2026-4201
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 12
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-09-542-552-6

Query Match 14.6%; Score 6; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EREKEQ 6
Db 1 EREKEQ 6

Search completed: August 16, 2004, 09:57:43
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:56:08 ; Search time 46 Seconds
(without alignments)
279.805 Million cell updates/sec

Title: US-09-856-070B-29
Perfect score: 41
Sequence: 1 EREKQMMREKEELMLRLQD.....EEKTKAERLSEIQIRALQ 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1292805 seqs, 313927144 residues

Word size : 5
Total number of hits satisfying chosen parameters: 9639

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/FCI_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/FCIUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	48.8	436	US-10-106-698-4728	Sequence 4728, Ap
2	20	48.8	586	US-10-116-275-117	Sequence 117, App
3	20	48.8	586	US-10-408-765A-2372	Sequence 2372, Ap
4	20	48.8	630	US-10-408-765A-1160	Sequence 1160, Ap
5	20	48.8	635	US-09-925-299-896	Sequence 896, App
6	20	48.8	635	US-09-925-299-896	Sequence 896, App
7	8	19.5	8	US-09-988-493-8	Sequence 8, Appli
8	8	19.5	105	US-10-424-599-232000	Sequence 232000,
9	8	19.5	170	US-10-425-114-43920	Sequence 43920, A
10	8	19.5	207	US-10-424-599-232005	Sequence 232005,
11	8	19.5	207	US-10-425-114-45766	Sequence 45766, A
12	8	19.5	288	US-10-424-598-232010	Sequence 232010,
13	8	19.5	391	US-10-424-599-232009	Sequence 232009,
14	8	19.5	408	US-10-425-114-55605	Sequence 55605, A
15	8	19.5	408	US-10-425-114-56199	Sequence 56199, A

16	7	17.1	12	US-10-424-599-229552	Sequence 229552,
17	7	17.1	9	US-09-848-294-5	Sequence 5, Appli
18	7	17.1	336	US-10-293-231-5	Sequence 5, Appli
19	7	17.1	458	US-10-108-260A-4328	Sequence 4328, Ap
20	7	17.1	499	US-10-087-192-597	Sequence 597, App
21	7	17.1	510	US-10-074-978A-34	Sequence 34, Appl
22	7	17.1	510	US-10-074-978A-224	Sequence 224, App
23	7	17.1	555	US-10-074-978A-36	Sequence 36, Appl
24	7	17.1	577	US-10-236-031B-54	Sequence 54, Appl
25	7	17.1	577	US-10-360-849A-27	Sequence 27, Appl
26	7	17.1	577	US-10-408-765A-453	Sequence 453, App
27	7	17.1	577	US-10-408-765A-454	Sequence 454, App
28	7	17.1	577	US-10-648-593-187	Sequence 187, App
29	7	17.1	583	US-10-205-219-151	Sequence 151, App
30	7	17.1	645	US-10-393-602-144	Sequence 144, App
31	7	17.1	650	US-10-104-047-3636	Sequence 3636, Ap
32	7	17.1	1215	US-09-817-913-11	Sequence 11, Appl
33	7	17.1	1215	US-09-376-280A-4	Sequence 4, Appli
34	7	17.1	1215	US-09-976-280A-22	Sequence 22, Appl
35	7	17.1	1215	US-09-817-538-11	Sequence 11, Appl
36	7	17.1	1215	US-09-800-187-6	Sequence 6, Appli
37	7	17.1	1215	US-10-438-075-2	Sequence 2, Appli
38	7	17.1	1215	US-10-360-534-6	Sequence 6, Appli
39	7	17.1	1312	US-10-393-602-148	Sequence 148, App
40	7	17.1	1318	US-10-408-765A-666	Sequence 666, App
41	7	17.1	1332	US-10-437-963-136245	Sequence 136245,
42	7	17.1	1377	US-10-437-963-136247	Sequence 136247,
43	7	17.1	1901	US-10-363-616-372	Sequence 372, App
44	6	14.6	15	US-09-933-780C-53	Sequence 53, Appl
45	6	14.6	45	US-10-424-599-226483	Sequence 226483,

ALIGNMENTS

RESULT 1

US-10-106-698-4728
; Sequence 4728, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8584
; SOFTWARE: Patent In Ver. 3.0
; SEQ ID NO 4728
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (382)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4728

Query Match 48.8%; Score 20; DB 14; Length 436;

Best Local Similarity 100.0%; Pred. No. 2.1e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQMMREKEELMLRLQD 20

DB 184 EREKQMMREKEELMLRLQD 203

```
RESULT 2
US-10-116-275-117
; Sequence 117, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elian Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Irela
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 117
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-117

Query Match      48.8%; Score 20; DB 15; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQMMREKEELMLRLQD 20
| | | | | | | | | | | | | | | | | | | | | |
Db 334 EREKQMMREKEELMLRLQD 353

RESULT 3
US-10-408-765A-2372
; Sequence 2372, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale B.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2372
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2372

Query Match      48.8%; Score 20; DB 15; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQMMREKEELMLRLQD 20
| | | | | | | | | | | | | | | | | | | | | |
Db 334 EREKQMMREKEELMLRLQD 353

RESULT 4
US-10-408-765A-1160
; Sequence 1160, Application US/10408765A
; Publication No. US20040101874A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale B.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1160
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1160

Query Match      48.8%; Score 20; DB 15; Length 630;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQMMREKEELMLRLQD 20
| | | | | | | | | | | | | | | | | | | | | |
Db 378 EREKQMMREKEELMLRLQD 397

RESULT 5
US-09-925-299-896
; Sequence 896, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 896
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-896

Query Match      48.8%; Score 20; DB 9; Length 635;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQMMREKEELMLRLQD 20
| | | | | | | | | | | | | | | | | | | | | |
Db 383 EREKQMMREKEELMLRLQD 402

RESULT 6
US-09-925-299-896
; Sequence 896, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
```

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 896
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-896

Query Match 48.8%; Score 20; DB 10; Length 635;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKEQMKREKELMLRLQD 20
Db 383 EREKEQMKREKELMLRLQD 402

RESULT 7
US-09-988-493-8
; Sequence 8, Application US/09988493
; Publication No. US2003006419A1
; GENERAL INFORMATION:
; APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri
; APPLICANT: O'Hare, Michael John
; APPLICANT: Page, Martin John
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Waterfield, Michael Derek
; TITLE OF INVENTION: Proteins, Genes, and Their Use for
; TITLE OF INVENTION: Diagnosis and Treatment of Breast Cancer
; FILE REFERENCE: 2543-1-024
; CURRENT APPLICATION NUMBER: US/09/988,493
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: PCT/GB01/01219
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: GB 0006695.1
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: GB 0007265.2
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: homo sapien
US-09-988-493-8

Query Match 19.5%; Score 8; DB 12; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 EKEELMLR 17
Db 1 EKEELMLR 8

RESULT 8
US-10-424-599-232000
; Sequence 232000, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 232000

; LENGTH: 105
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51518C.1.pep
US-10-424-599-232000

Query Match 19.5%; Score 8; DB 12; Length 105;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TKKAEREL 32
Db 64 TKKAEREL 71

RESULT 9
US-10-425-114-43920
; Sequence 43920, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43920
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700728590_FLI.pep
US-10-425-114-43920

Query Match 19.5%; Score 8; DB 12; Length 170;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TKKAEREL 32
Db 129 TKKAEREL 136

RESULT 10
US-10-424-599-232005
; Sequence 232005, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 232005
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51522C.1.pep
US-10-424-599-232005

Query Match 19.5%; Score 8; DB 12; Length 207;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TKKAEREL 32
| | | | | | | |
Db 166 TKKAEREL 173

RESULT 11

US-10-425-114-45766
; Sequence 45766, Application US/10425114
; Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 45766

LENGTH: 207

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: 700739674_FLI.pep

US-10-425-114-45766

Query Match 19.5%; Score 8; DB 12; Length 207;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TKKAEREL 32
| | | | | | | |
Db 166 TKKAEREL 173

RESULT 12

US-10-424-599-232010
; Sequence 232010, Application US/10424599
; Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 232010

LENGTH: 298

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

NAME/KEY: unsure

LOCATION: (1) (298)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_51527C.1.pep

US-10-424-599-232010

Query Match 19.5%; Score 8; DB 12; Length 298;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TKKAEREL 32
| | | | | | | |
Db 257 TKKAEREL 264

RESULT 13

US-10-424-599-232009
; Sequence 232009, Application US/10424599
; Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 232009

LENGTH: 391

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_51526C.1.pep

US-10-424-599-232009

Query Match 19.5%; Score 8; DB 12; Length 391;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TKKAEREL 32
| | | | | | | |
Db 351 TKKAEREL 358

RESULT 14

US-10-425-114-55605
; Sequence 55605, Application US/10425114
; Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 55605

LENGTH: 408

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: UC-GMCRONIFR018B10_FLI.pep

US-10-425-114-55605

Query Match 19.5%; Score 8; DB 12; Length 408;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TKKAEREL 32
| | | | | | | |
Db 368 TKKAEREL 375

RESULT 15

Mon Aug 16 10:46:50 2004

```

US-10-425-114-56199
; Sequence 56199, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313) B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56199
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMROPIC090E09_FLI.pep
US-10-425-114-56199

```

```

Query Match      19.5%; Score 8; DB 12; Length 408;
Best Local Similarity 100.0%; Pred.No. 9.7; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

```

```

QY      25 TKKAEREL 32
Db      368 TKKAEREL 375

```

Search completed: August 16, 2004, 10:02:25
Job time : 47 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:53:32 ; Search time 38 Seconds
(without alignments)
103.786 Million cell updates/sec

Title: US-09-856-070b-29

Perfect score: 41

Sequence: 1 EREKEQMMREKEELMLRLQD.....BEKTKAERELSEIQIRALQ 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 5
Total number of hits satisfying chosen parameters: 3110

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR 78:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	20	48.8	581	2 I45889	ezrin - bovine
2	20	48.8	586	1 A34400	ezrin [validated]
3	20	48.8	630	2 T47177	hypothetical prote
4	12	29.3	586	1 B41129	ezrin - mouse
5	7	17.1	142	2 C86389	17.3K hypothetical
6	7	17.1	239	2 H64864	DNA-binding protei
7	7	17.1	239	2 AF0261	fatty acid metabol
8	7	17.1	239	2 B85897	DNA-binding protei
9	7	17.1	239	2 B98839	DNA-binding protei
10	7	17.1	239	2 AF0723	fatty acid-fatty a
11	7	17.1	279	2 B82144	fatty acid metabol
12	7	17.1	303	2 F90478	epimerase, probabl
13	7	17.1	311	2 G90407	homoserine kinase
14	7	17.1	452	2 T40769	hypothetical prote
15	7	17.1	583	2 S13786	DNA-directed DNA p
16	7	17.1	577	1 A41289	moesin - human
17	7	17.1	577	1 S39804	moesin - pig
18	7	17.1	583	1 A46127	radixin - human
19	7	17.1	583	1 S39805	radixin - pig
20	7	17.1	583	1 A41129	radixin - mouse
21	7	17.1	612	2 T00384	hypothetical prote
22	7	17.1	668	2 H71312	probable ATP-depen
23	7	17.1	1051	2 T51904	hypothetical prote
24	7	17.1	1095	2 T13964	probable histone d
25	7	17.1	1120	2 T38431	DNA-directed RNA p
26	7	17.1	1129	2 T42732	A-kinase anchoring
27	7	17.1	1308	2 T05178	hypothetical prote
28	7	17.1	1312	2 T30845	probable DNA repai
29	6	14.6	47	2 A71340	hypothetical prote

30	6	14.6	112	2	H82197	transcription repr
31	6	14.6	112	2	B96951	similar to arsenat
32	6	14.6	116	2	F90024	hypothetical prote
33	6	14.6	124	2	A70402	flagellar basal bo
34	6	14.6	139	2	B65036	robable thioredoxi
35	6	14.6	139	2	G85904	probable thioredox
36	6	14.6	139	2	H91059	probable thioredox
37	6	14.6	139	2	AB0831	thioredoxin 2 [imp
38	6	14.6	145	2	F69160	hypothetical prote
39	6	14.6	151	2	T11571	hypothetical prote
40	6	14.6	163	1	S01399	H+-transporting tw
41	6	14.6	166	2	S22631	petP protein - Rho
42	6	14.6	169	2	D88040	protein F47F6.4 [i
43	6	14.6	186	2	G71725	ribosome recycling
44	6	14.6	190	2	G70239	conserved hypothe
45	6	14.6	191	2	S43178	gene sj22 protein

ALIGNMENTS

RESULT 1

I45889
ezrin - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 24-Nov-2003
C:Accession: I45889
R:Berenson, C.M.; Zhao, H.; Saitoh, K.; Duman, R.S.; Nestler, E.J.
Mol. Cell. Neurosci. 4, 64-73, 1993
A:Title: Ezrin and osteonectin, two proteins associated with cell shape and growth, are
A:Reference number: I45889
A:Accession: I45889
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-581 <B8R>
A:Cross-references: GB:M98498; NID:G289407; PID:AAA30510.1; PID:G289408
C:Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology
F,7-291/Domain: protein 4.1 membrane-binding domain homology <B41>

Query Match 48.8%; Score 20; DB 2; Length 581;
Best Local Similarity 100.0%; Pred. No. 2.8e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

QY 1 EREKEQMMREKEELMLRLQD 20
DB 334 EREKEQMMREKEELMLRLQD 353
|||||

RESULT 2

A34400
ezrin [validated] - human
N:Alternate names: cytoovillin; p81 protein; villin 2
C:Species: Homo sapiens (man)
C:Date: 22-Jun-1990 #sequence_revision 14-Jul-1994 #text_change 24-Nov-2003
C:Accession: A34400; S09263; E61002
R:Turunen, O.; Winkvist, R.; Pakkanen, R.; Grzeschik, K.H.; Wahlestrom, T.; Vaheri, A.
J. Biol. Chem. 264, 16727-16732, 1989
A:Title: Cytoovillin, a microvillar M-r 75,000 protein. cDNA sequence, prokaryotic expres
A:Reference number: A34400; MUID:89380299; PMID:2674140
A:Accession: A34400
A:Molecule type: mRNA
A:Residues: 1-586 <IUR>
A:Cross-references: GB:J05021
A:Note: the translation of residues 1-11 is not given
A:Note: parts of this sequence were confirmed by protein sequencing
R:Gould, K.L.; Bretscher, A.; Esch, F.S.; Hunter, T.
EMBO J. 8, 4133-4142, 1989
A:Title: cDNA cloning and sequencing of the protein-tyrosine kinase substrate, ezrin, re
A:Reference number: S09263; MUID:90076135; PMID:2591371
A:Accession: S09263
A:Molecule type: mRNA
A:Residues: 2-586 <GOU>
A:Cross-references: GB:X51521; NID:g31282; PIDN:CAA35893.1; PID:g31283

R;Bauw, G.; Rasmussen, H.H.; Van Den Bulcke, M.; Van Damme, J.; Puype, M.; Gesser, B.; C
Electrophoresis 11, 528-536, 1990
A;Title: Two-dimensional gel electrophoresis, protein electrophoresis and microsequencing
A;Reference number: A61002; MUID:91031404; PMID:1699755
A;Accession: E61002

A;Molecule type: protein
A;Residues: 255-263;194,'Q',196-199,'X',201,264-270 <BAU>
A;Note: it is not certain whether this material represents ezrin or radixin (see entry A
A;Note: this material corresponds to transformed epithelial amion cell (AVA) database
C;Comment: This protein is located in microvilli and is proposed to play a role in modul
C;Genetics:

A;Gene: GDB:VIL2
A;Cross-references: GDB:120489; OMIM:123900
A;Map position: 6q25-6q26
C;Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology
C;Keywords: actin binding; cytoskeleton; membrane-associated protein; phosphoprotein
F;2-586/Product: ezrin #status experimental <MAT>
F;7-291/Domains: protein 4.1 membrane-binding domain homology <B41>
F;553-586/Region: actin binding #status predicted
F;66/Binding site: phosphate (Ser) (covalent) #status predicted
F;214,299,332/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 48.8%; Score 20; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.8e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQMPREKELMLRLQD 20
|||||
DB 334 EREKQMPREKELMLRLQD 353

RESULT 3
T47177
Hypothetical protein DKFZp762H157.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 24-Nov-2003
C;Accession: T47177
R;Ottewaelde, B.; Obermaier, B.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24377
A;Accession: T47177
A;Molecule type: mRNA
A;Residues: 1-630 <AAA>
A;Cross-references: EMBL:AL162086
A;Experimental source: adult melanoma (Mewo cell line); clone DKFZp762H157
C;Genetics:

A;Note: DKFZp762H157.1
C;Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology

Query Match 48.8%; Score 20; DB 2; Length 630;
Best Local Similarity 100.0%; Pred. No. 3e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREKQMPREKELMLRLQD 20
|||||
DB 378 EREKQMPREKELMLRLQD 397

RESULT 4
B41129
ezrin - mouse
N;Alternate names: cyto villin; p81 protein; radixin; villin 2
C;Species: Mus musculus (house mouse)
C;Date: 03-Aug-1992 #sequence_revision 14-Jul-1994 #text_change 24-Nov-2003
C;Accession: B41129; C46501; A46501; B46501; S24200
R;Funayama, N.; Nagafuchi, A.; Sato, N.; Tsukita, S.; Tsukita, S.
J. Cell Biol. 115, 1039-1048, 1991
A;Title: Radixin is a novel member of the band 4.1 family.
A;Reference number: A41129; MUID:92064635; PMID:1955455
A;Accession: B41129
A;Molecule type: mRNA
A;Residues: 1-586 <FUN>

A;Cross-references: EMBL:X60671; NID:950880; PIDN:CAA43086.1; PID:950881
R;Egerton, M.; Burgess, W.H.; Chen, D.; Drucker, B.J.; Bretscher, A.; Samelson, L.E.
J. Immunol. 149, 1847-1852, 1992
A;Title: Identification of ezrin as an 81-kDa tyrosine-phosphorylated protein in T cells.
A;Reference number: A46501; MUID:92388649; PMID:1381389
A;Accession: C46501
A;Status: preliminary
A;Molecule type: protein
A;Residues: 412-426 <EGE>
A;Experimental source: MRL lpr/lpr, T-cells
A;Note: sequence extracted from NCBI backbone (NCBIP:112938)
A;Accession: A46501
A;Status: preliminary
A;Molecule type: protein
A;Residues: 27-33,'E' <EG2>
A;Experimental source: MRL lpr/lpr, T-cells
A;Note: sequence extracted from NCBI backbone (NCBIP:112936)
A;Accession: B46501
A;Status: preliminary
A;Molecule type: protein
A;Residues: 53-57,148,'L',150,'G',152-155 <EG3>
A;Experimental source: MRL lpr/lpr, T-cells
A;Note: sequence extracted from NCBI backbone (NCBIP:112940)
C;Comment: This protein is located in microvilli and is proposed to play a role in modula
C;Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology
C;Keywords: actin binding; cytoskeleton; cytosol; membrane-associated protein; phosphop
F;2-586/Product: ezrin #status predicted <MAT>
F;7-291/Domains: protein 4.1 membrane-binding domain homology <B41>
F;553-586/Region: actin binding #status predicted
F;66/Binding site: phosphate (Ser) (covalent) #status predicted
F;214,299,332/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 29.3%; Score 12; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 0.00026; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 REKEELMLRLQD 20
|||||
DB 342 REKEELMLRLQD 353

RESULT 5
C86389
17.3K hypothetical protein F28923.5 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001
C;Accession: C86389
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86389
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-142 <STO>
A;Cross-references: GB:AE005172; NID:g11079495; PIDN:AG29206.1; GSPDB:GNC0141
C;Genetics:
A;Map position: 1

Query Match 17.1%; Score 7; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 7.7; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 ETKKAE 29
|||||
DB 102 ETKKAE 108

A:Molecule type: DNA
 A:Residues: 1-239 <KUR>
 A:Cross-references: GB:ALS90842; PIDN:CAK90954.1; PID:gl5980150; GSPDB:GN00175
 C:Genetics:
 A:Gene: fadR
 C:Superfamily: pyruvate dehydrogenase complex repressor

Query Match 17.1%; Score 7; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Gaps 0;

Qy 28 AERESE 34
 Db 33 AERESE 39

RESULT 8
 B85697
 DNA-binding protein, fatty acid/fatty acyl-responsive - Escherichia coli (strain O157:H7)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 19-Jul-2002
 C:Accession: B85697
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Poramoussis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: B85697
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-239 <STC>
 A:Cross-references: GB:AE005174; NID:gl2514884; PIDN:AAG56038.1; GSPDB:GN00145; UWGP:Z19:
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: fadR
 C:Superfamily: pyruvate dehydrogenase complex repressor

Query Match 17.1%; Score 7; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AERESE 34
 Db 33 AERESE 39

RESULT 9
 B9839
 DNA-binding protein, fatty acid/fatty acyl-responsive - Escherichia coli (strain O157:H7)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 19-Jul-2002
 C:Accession: B9839
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genon
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: B9839
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-239 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA835105.1; PID:gl3361146; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs1682
 C:Superfamily: pyruvate dehydrogenase complex repressor

Query Match 17.1%; Score 7; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AERESE 34
 Db 33 AERESE 39

RESULT 10
 AF0261
 Fatty acid metabolism regulatory protein fadR [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 19-Jul-2002
 C:Accession: AF0261
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AF0261
 A:Status: preliminary

Query Match 17.1%; Score 7; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AERESE 34
 Db 33 AERESE 39

RESULT 11
 AF0261
 Fatty acid metabolism regulatory protein fadR [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 19-Jul-2002
 C:Accession: AF0261
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AF0261
 A:Status: preliminary

Query Match 17.1%; Score 7; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AERESE 34
 Db 33 AERESE 39

RESULT 12
 AF0261
 Fatty acid metabolism regulatory protein fadR [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 19-Jul-2002
 C:Accession: AF0261
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AF0261
 A:Status: preliminary

Query Match 17.1%; Score 7; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AERESE 34
 Db 33 AERESE 39

RESULT 13
 AF0261
 Fatty acid metabolism regulatory protein fadR [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 19-Jul-2002
 C:Accession: AF0261
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AF0261
 A:Status: preliminary

Query Match 17.1%; Score 7; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AERESE 34
 Db 33 AERESE 39

RESULT 14
 AF0261
 Fatty acid metabolism regulatory protein fadR [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 19-Jul-2002
 C:Accession: AF0261
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AF0261
 A:Status: preliminary

Query Match 17.1%; Score 7; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AERESE 34
 Db 33 AERESE 39

RESULT 15
 AF0261
 Fatty acid metabolism regulatory protein fadR [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 19-Jul-2002
 C:Accession: AF0261
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession

Db 33 AERELSE 39

RESULT 10

AF0723
fatty acid-fatty acyl responsive DNA-binding protein [imported] - Salmonella enterica subsp. enterica serovar typhi
C;Species: Salmonella enterica subsp. enterica serovar typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AF0723
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulle, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar typhi
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AF0723
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-239 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05488.1; PID:gl6502992; GSPDB:GN00176
C;Genetics:
A;Gene: STY1934
C;Superfamily: pyruvate dehydrogenase complex repressor

Query Match 17.1%; Score 7; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AERELSE 34

|||||

Db 33 AERELSE 39

RESULT 11

B82144
fatty acid metabolism regulator protein VC1900 [imported] - Vibrio cholerae (strain N169)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 19-Jul-2002
C;Accession: B82144
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Church, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: B82144
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-279 <HEI>
A;Cross-references: GB:AE004255; GB:AE003852; NID:9656424; PIDN:AAF95048.1; GSPDB:GN00176
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1900
A;Map position: 1
C;Superfamily: pyruvate dehydrogenase complex repressor

Query Match 17.1%; Score 7; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AERELSE 34

|||||

Db 33 AERELSE 39

RESULT 12

F90478
epimerase, probable [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: F90478

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-y Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.; arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: F90478
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-303 <KUR>
A;Cross-references: GB:AE006641; NID:gl3816366; PIDN:AAK43085.1; GSPDB:GN00155
C;Genetics:
A;Gene: SS02981

Query Match 17.1%; Score 7; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 KKAEREL 32

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

|||||

Db 275 KKAEREL 281

Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 KXAEREL 32

Db 122 KXAEREL 128

Db 315 RELSEQUI 321

Search completed: August 16, 2004, 09:56:52
Job time : 38 secs

RESULT 15

S13786
DNA-directed DNA polymerase (EC 2.7.7.7) III chain dnaX - Bacillus subtilis
N/Alternate names: DNA polymerase III (gamma and tau subunits) dnaX
C/Species: Bacillus subtilis
C/Date: 19-May-1994 #sequence revision 19-May-1994 #text_change 20-Jun-2000
C/Accession: S13786; S00745; S66049; B69618
R/Alonso, J.C.; Shirahige, K.; Ogasawara, N.
Nucleic Acids Res. 18, 6771-6777, 1990
A/Title: Molecular cloning, genetic characterization and DNA sequence analysis of the re
A/Reference number: S13786; MUID:91088245; PMID:2124672
A/Accession: S13786
A/Molecule type: DNA
A/Residues: 1-563 <ALO>
A/Cross-references: EMBL:X17014; NID:9453238; PIDN:CAA34877.1; PID:g580914
R/Struck, J.C.R.; Vogel, D.W.; Ubrich, N.; Erdmann, V.A.
Nucleic Acids Res. 16, 2720, 1988
A/Title: A dnaX-like open reading frame downstream from the Bacillus subtilis scRNA gen
A/Reference number: S00745; MUID:88203213; PMID:2452406
A/Accession: S00745
A/Molecule type: DNA
A/Residues: 1-422 <STR>
A/Cross-references: EMBL:X06803; NID:g39891; PIDN:CAA29958.1; PID:g580855
R/Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA Res. 1, 1-14, 1994
A/Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chrom
A/Reference number: S65967; MUID:96051385; PMID:7584024
A/Accession: S66049
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-563 <OGA>
A/Cross-references: EMBL:D26185; NID:g467326; PIDN:BAA05255.1; PID:g467409
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Buschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: B69618
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-563 <KUN>
A/Cross-references: GB:Z99104; GB:AL009126; NID:g2632267; PIDN:CAB11795.1; PID:g2632286
A/Experimental source: strain 168
C/Genetics:
A/Gene: dnaX; dnaZX
A/Start codon: GTG
C/Superfamily: DNA-directed DNA polymerase III gamma chain
C/Keywords: nucleotidyltransferase

Query Match 17.1%; Score 7; DB 2; Length 563;

Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 RELSEQUI 36

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 09:52:02 ; Search time 32 Seconds
(without alignments)
66.715 Million cell updates/sec

Title: US-09-856-070B-29
Perfect score: 41
Sequence: 1 BREKQMMREKEELMLRLQD.....BEKTKAERLSEIQIRALQ 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 5

Total number of hits satisfying chosen parameters: 1807

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42.*

Result No.	Score	Query Match	Length	DB ID	Description
1	20	48.8	580	1	EZRI_BOVIN
2	20	48.8	585	1	EZRI_HUMAN
3	12	29.3	585	1	EZRI_MOUSE
4	7	17.1	238	1	PADR_ECOLI
5	7	17.1	238	1	PADR_SALTI
6	7	17.1	238	1	PADR_SALTY
7	7	17.1	238	1	PADR_YERPE
8	7	17.1	240	1	PADR_SHEON
9	7	17.1	279	1	PADR_VIBCH
10	7	17.1	279	1	PADR_VIBPA
11	7	17.1	279	1	PADR_VIBVU
12	7	17.1	311	1	KHSE_SULSO
13	7	17.1	563	1	DP3X_BACSU
14	7	17.1	576	1	MOES_HUMAN
15	7	17.1	576	1	MOES_MOUSE
16	7	17.1	576	1	MOES_PIG
17	7	17.1	576	1	MOES_RAT
18	7	17.1	583	1	RADI_CHICK
19	7	17.1	583	1	RADI_HUMAN
20	7	17.1	583	1	RADI_MOUSE
21	7	17.1	583	1	RADI_PIG
22	7	17.1	585	1	EZRI_RABIT
23	7	17.1	1005	1	EVC_MOUSE
24	7	17.1	1120	1	RPOM_SCHPO
25	7	17.1	1129	1	AK11_RAT
26	7	17.1	1149	1	HDA6_MOUSE
27	7	17.1	1215	1	HDA6_HUMAN
28	7	17.1	1901	1	AK11_HUMAN
29	6	14.6	47	1	X311_TREPA
30	6	14.6	95	1	CH10_STEMU
31	6	14.6	103	1	CYT5_PIG
32	6	14.6	103	1	CIT8_PIG
33	6	14.6	112	1	RSRI_VIBCH

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1	34	6	14.6	124	1	FLGB_AQUAE	067244 aquifex aeo
EZRI_BOVIN	35	6	14.6	139	1	THI2_ECOLI	P3636 escherichia
ID_EZRI_BOVIN	36	6	14.6	152	1	Y972_THETN	Q8rb52 thermoanaer
AC P31976	37	6	14.6	168	1	ATPF_BACCA	P41014 bacillus ca
DT 01-JUL-1993 (Rel. 26, Last sequence update)	38	6	14.6	162	1	VB48_METTW	Q50521 methanobact
DT 01-JUL-1993 (Rel. 26, Last sequence update)	39	6	14.6	162	1	ATPF_BACPP3	P09221 bacillus ps
DT 10-OCT-2003 (Rel. 42, Last annotation update)	40	6	14.6	186	1	PETP_RHOCA	P31078 rhodobacter
DE Ezrin (p81) (Cytovillin) (Villin 2)	41	6	14.6	170	1	DEF1_BORBR	Q7wg99 bordetella
GN VII2	42	6	14.6	170	1	DEF1_BORPA	Q7wv88 bordetella
OS Bos taurus (Bovine)	43	6	14.6	170	1	DEF2_BORPE	Q9ze08 rickettsia
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	44	6	14.6	186	1	RRF_RICPR	P77181 escherichia
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	45	6	14.6	196	1	PAAY_ECOLI	
OC Bovidae; Bovinae; Bos.							
OX NCBI_TaxID=9913;							
RP SEQUENCE FROM N.A.							
RC TISSUE=Brain;							
RA Bergson C.M., Zhao H., Saijoh K., Duman R.S., Nestler E.J.;							
RT "Ezrin and osteonectin, two proteins associated with cell shape and							
RT growth, are enriched in the locus coeruleus.";							
RL Mol. Cell. Neurosci. 4:64-73(1993).							
RN [2]							
RP SEQUENCE OF 1-15 AND 126-140.							
RC TISSUE=Kidney;							
RX MEDLINE=96239137; PubMed=8660651;							
RA Galat A., Gerbod M.C., Bouet F., Riviere S.;							
RT "Proteins and their amino acid compositions: uniqueness, variability,							
RT and applications.";							
RL Arch. Biochem. Biophys. 330:229-237(1996).							
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal							
CC structures to the plasma membrane.							
CC -!- SUBCELLULAR LOCATION: Microvillar peripheral membrane protein							
CC (cytoplasmic side).							
CC -!- PTM: Phosphorylated by tyrosine-protein kinases.							
CC -!- SIMILARITY: Contains 1 FERM domain.							
CC -----							
CC This SWISS-PROT entry is copyright. It is produced through a collaboration							
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -							
CC the European Bioinformatics Institute. There are no restrictions on its							
CC use by non-profit institutions as long as its content is in no way							
CC modified and this statement is not removed. Usage by and for commercial							
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/							
CC or send an email to license@isb-sib.ch).							
CC -----							
CC EMBL; M98498; AAA30510.1; -							
DR PIR; I45889; I45889.							
DR InterPro; IPR000299; Band_4.1.							
DR InterPro; IPR000798; Ez/rad/moesin.							
DR InterPro; IPR008954; Moesin.							
DR Pfam; PF00373; Band_41; 1.							
DR Pfam; PF00769; ERM; 1.							
DR PRINTS; PR00935; BAND41.							
DR SMART; SM00295; B41; 1.							
DR PROSITE; PS00660; FERM_1; 1.							
DR PROSITE; PS00661; FERM_2; 1.							
DR PROSITE; PS50057; FERM_3; 1.							

Best Local Similarity 100.0%; Pred. No. 2.8e-12; Mismatches 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 BREKEQWREKEELMLRLQD 20
 |||||
 DB 333 BREKEQWREKEELMLRLQD 352
 |||||

RESULT 3
 EZRI MOUSE
 ID EZRI MOUSE STANDARD; PRT; 585 AA.
 AC P26040; Q80ZTS; Q9DC11;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ezrin (p81) (Cytovillin) (Willin 2).
 GN VIL2.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92064635; PubMed=1955455;
 RA Funayama N., Nagafuchi A., Sato N., Tsukita S., Tsukita S.;
 RT "Radixin is a novel member of the band 4.1 family.";
 RL J. Cell Biol. 115:1039-1048(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Kidney;
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito K., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Kasagawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragan T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kongaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wählstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang N.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shingawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.P., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16893-16903(2002).
 CC -I- FUNCTION: Probably involved in connections of major cytoskeletal
 CC structures to the plasma membrane.
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -I- TISSUE SPECIFICITY: Component of the microvilli of intestinal
 CC epithelial cells.
 CC -I- FTM: Phosphorylated by tyrosine-protein kinases.
 CC -I- SIMILARITY: Contains 1 FERM domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X60671; CAA43086.1; -;
 CC EMBL; AC002766; BAB22341.1; -;
 CC EMBL; BC048181; AAH48181.2; -;
 CC FIRM; B41129; B41129.
 CC MGD; MGI:98931; Vll2.
 CC InterPro; IPR000299; Band 4.1.
 CC InterPro; IPR000798; Ez/rad/moesin.
 CC Pfam; PF00373; Band 41; 1.
 CC Pfam; PF00769; ERM; 1.
 CC PRINTS; PR00935; BAND41.
 CC SMART; SM00235; B41; 1.
 CC PROSITE; PS00660; FERM_1; 1.
 CC PROSITE; PS00661; FERM_2; 1.
 CC PROSITE; PS00557; FERM_3; 1.
 CC Structural protein; Cytoskeleton; Phosphorylation.
 CC BY SIMILARITY.
 KW INIT MET 0
 FT DOMAIN 1 294
 FT MOD_RES 145 145
 FT MOD_RES 353 353
 FT MOD_RES 47 47
 FT CONFLICT 324 324 Q -> P (IN REF. 1).
 FT CONFLICT 569 569 T -> A (IN REF. 1).
 FT CONFLICT 569 569 Q -> R (IN REF. 2).
 SQ SEQUENCE 585 AA; 69275 MW; 205CED2E326A0562 CRC64;
 Query Match 29.3%; Score 12; DB 1; Length 585;
 Best Local Similarity 100.0%; Pred. No. 0.00019;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 REKEELMLRLQD 20
 |||||
 DB 341 REKEELMLRLQD 352
 |||||

RESULT 4
 PADR_ECOLI
 ID PADR_ECOLI STANDARD; PRT; 238 AA.
 AC P09371; P76827;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Fatty acid metabolism regulator protein.
 GN PADR OR OLER OR THDB OR B1187 OR C1635 OR Z1950 OR ECS1682 OR SF1176
 OR S1284.
 OS Escherichia coli,

OS Escherichia coli O6,
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562, 217992, 83334, 623;
RN [1]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=K12;
RC MEDLINE=88335542; PubMed=2843809;
RA Dirusso C.C.;
RT "Nucleotide sequence of the fadR gene, a multifunctional regulator of
RT fatty acid metabolism in Escherichia coli";
RL Nucleic Acids Res. 16:7995-8009(1988).
RN [2]
RN REVISION TO 146.
RP SPECIES=E.coli; STRAIN=K12;
RC Dirusso C.C.;
RA Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:12453-1474 (1997).
RN [4]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=K12;
RC MEDLINE=97061202; PubMed=8905232;
RX Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [5]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RC MEDLINE=22388234; PubMed=12471157;
RX Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.F., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN [6]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [7]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=O157:H7 / RMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Yokoyama K.,
RA Han C.-G., Ohsuho E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [8]
RN PARTIAL SEQUENCE, AND CHARACTERIZATION.
RP SPECIES=E.coli;
RC MEDLINE=92235103; PubMed=1569108;
RX Dirusso C.C.; Heimert T.L., Metzger A.K.;
RA "Characterization of FadR, a global transcriptional regulator of
RA fatty acid metabolism in Escherichia coli. Interaction with the fadB
RA promoter is prevented by long chain fatty acyl coenzyme A.";
RL J. Biol. Chem. 267:8685-8691(1992).
RN [9]
RN SEQUENCE FROM N.A.
RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RC MEDLINE=2272406; PubMed=12394590;
RX Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [10]
RN SEQUENCE FROM N.A.
RP SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RC MEDLINE=22590274; PubMed=12704152; Venkatesan M.M., Deng W.,
RA Wei C., Goldberg M.B., Burland V., Plunkett G. III, Rose D.J., Darling A.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Runyen-Janecky L.J., Zhou S.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T";
RL Infect. Immun. 71:2775-2786(2003).
RN [11]
RN CHARACTERIZATION.
RP SPECIES=E.coli;
RC MEDLINE=98049521; PubMed=9388199;
RX Raman N., Black P.N., Dirusso C.C.;
RA "Characterization of the fatty acid-responsive transcription factor
RA FadR. Biochemical and genetic analyses of the native conformation and
RA functional domains.";
RL J. Biol. Chem. 272:30645-30650(1997).
RN [12]
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RP MEDLINE=21192039; PubMed=11296236;
RX van Aalten D.M., Dirusso C.C., Knudsen J.;
RA "The structural basis of acyl coenzyme A-dependent regulation of the
RA transcription factor FadR.";
RL EMBO J. 20:2041-2050(2001).
RN [13]
RN X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RP MEDLINE=21264500; PubMed=11279025;
RX Xu Y., Heath R.J., Li Z., Rock C.O., White S.W.;
RA "The FadR-DNA complex. Transcriptional control of fatty acid
RA metabolism in Escherichia coli";
RL J. Biol. Chem. 276:17373-17379(2001).
RN [14]
RN FUNCTION: Multifunctional regulator of fatty acid metabolism.
CC Represses transcription of at least eight genes required for fatty
CC acid transport and beta-oxidation among which fadA, fadB, fadD,
CC fadL and fadE. Activates transcription of at least three genes
CC required for unsaturated fatty acid biosynthesis: fabA, fabB and
CC iclR, the gene encoding the transcriptional regulator of the
CC aceBAK operon encoding the glyoxylate shunt enzymes. Binding of
CC fadR is specifically inhibited by long chain fatty acyl-CoA
CC compounds.
CC [1] SUBUNIT: Homodimer.
CC [2] SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC [3] DOMAIN: The C-terminal domain bind acyl-CoA.
CC [4] SIMILARITY: Contains 1 HTH gntR-type DNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

-----
EMBL; X08087; CAA30881.1; -.
EMBL; AE000217; AAC74271.1; -.
EMBL; D90753; BAA36042.1; -.
EMBL; D90752; BAA36034.1; -.
EMBL; AE016759; AAN80100.1; -.
EMBL; AE053335; AAG56038.1; -.
EMBL; AF002555; BAB35105.1; -.
EMBL; AE015144; AAN42791.1; ALT_INIT.
EMBL; AE016982; AAP16682.1; -.
PIR; B85697; B85697.
PIR; B99839; B99839.
PIR; H64864; H64864.
PDB; 1H9G; 15-APR-01.
PDB; 1HW1; 06-JUN-01.
PDB; 1HW2; 04-JUN-01.
EcoGene; EG10281; fadr.
HAMAP; MF_00696; -.
InterPro; IPR008920; Fadr_C.
InterPro; IPR000524; HTH_GntR.
Pfam; PF00392; gntR; 1.
PRINTS; PR00035; HTHGNTR.
SMART; SM00345; HTH_GNTR; 1.
PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
Fatty acid metabolism; Transcription regulation; Activator; Repressor;
DNA-binding; Complete proteome; 3D-structure.
INIT_MET 0 0
FT DOMAIN 1 70 HTH GNTR-TYPE.
FT DNA_BIND 33 52 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 238 AA; 26837 MW; ADA5806B7633197E CRC64;

```

```
Nature 413:848-852(2001)).
```

```
[2]
RRL SEQUENCE FROM N.A.
```

```
RNP STRAIN=TY2 / ATCC 700931;
RRP MEDLINE=22531367; PubMed=12644504;
RCX Deng W., Liou S.-R., Plunkett G.III, Mayhew G.F., Rose D.J.,
RCY Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RRA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RRT and Cti18";
RRI J. Bacteriol. 185:2330-2337(2003).
```

```
RRL CC -!- FUNCTION: Multifunctional regulator of fatty acid metabolism.
CCC Represses transcription of at least eight genes required for fatty
CCC acid transport and beta-oxidation among which fadA, fadB, fadD,
CCC fadI and fadE. Activates transcriptions of at least three genes
CCC required for unsaturated fatty acid biosynthesis: fabA, fabB and
CCC iclR, the gene encoding the transcriptional regulator of the
CCC acetoacetyl operon encoding the glyoxylate shunt enzymes. Binding of
CCC fadR is specifically inhibited by long chain fatty acyl-CoA
CCC compounds (By similarity).
CCC -!- SUBUNIT: Homodimer (By similarity).
```

```
CCC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CCC -!- SIMILARITY: Contains 1 HTH GNTR-type DNA-binding domain.
-----
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)
or send an email to licensee@isb-sib.ch) .
-----
```

```
EMBL; AL627272; CAD05488.1; -.
DR EMBO; AE016837; RAO68738.1; -.
DR HAMAP; MF_00696; -.
DR InterPro; IPR008920; Fadr_C.
DR InterPro; IPR000524; HTH_Gntr.
DR Pfam; PF00392; gnrcr; 1.
DR PRINTS; PR00035; HTHGNTR.
DR SMART; SM00345; HTH_GNTR; 1.
DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
DR Fatty acid metabolism; Transcription regulation; Activator; Repressor;
```

```
KW DNA-binding; Complete proteome.
KW INIT_MET 0 BY SIMILARITY.
FT DOMAIN 1 70 HTH GNTR-TYPE.
FT DNA_BIND 33 52 H-T-H MOTIF (POTENTIAL)..
SQ SEQUENCE 238 AA; 2680 MW; 9E7EB7DE6F72CB7 CRC64;
```

```
Query Match 17.1%; Score 7; DB 1; Length 238;
Best Local Similarity 100.0%; Pred.No. 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY QY 28 AERESE 34
Db Db |||||
Db Db 32 AERESE 38
```

```
RESULT 6
ID_FADR_SALTY STANDARD; PRT; 238 AA.
AC O8ZPI5;
DT 28-FEB-2003 (Rel. 41, Created)
```

```
DT 28-FEB-2003 (Rel. 41, Last sequence update)
```

```
DT 10-OCT-2003 (Rel. 42, Last annotation update)
```

```
DE DE Fatty acid metabolism regulator protein.
```

```
GK FADR OR STM1805.
```

```
OS Salmonella typhimurium.
```

```
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales;
```

```
OC Enterobacteriaceae; Salmonella.
```

```
OX NCBI_TaxID=602;
```

```
[1]
```

```
SEQUENCE FROM N.A.
STRAIN=Lf2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
```

McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Floreal L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
 "Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2.";
 Nature 413:852-856(2001).
 RL Nature 413:852-856(2001).
 CC -1- FUNCTION: Multifunctional regulator of fatty acid metabolism.
 CC Represses transcription of at least eight genes required for fatty
 CC acid transport and beta-oxidation among which fadA, fadB, fadD,
 CC fadL and fadE. Activates transcription of at least three genes
 CC required for unsaturated fatty acid biosynthesis: fabA, fabB and
 CC iclR, the gene encoding the transcriptional regulator of the
 CC aceBAK operon encoding the glyoxylate shunt enzymes. Binding of
 CC fadR is specifically inhibited by long chain fatty acyl-CoA
 CC compounds (By similarity).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: Contains 1 HTH gntR-type DNA-binding domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AB008780; AAL20720.1; -
 CC STYGene: SG2222; fadR.
 CC HAMAP: MF_00696; -; 1.
 CC InterPro: IPR008920; FadR_C.
 CC InterPro: IPR000524; HTH_GntR.
 CC Pfam: PF00392; gntR; 1.
 CC PRINTS: PR00035; HTHGNTR.
 CC SMART: SM00345; HTH_GNTR; 1.
 CC PROSITE: PS00043; HTH_GNTR_FAMILY; 1.
 CC Fatty acid metabolism; Transcription regulation; Activator; Repressor;
 KW DNA-binding; Complete proteome.
 FT INIT MET 0 0 BY SIMILARITY.
 FT DOMAIN 1 70 HTH_GNTR-TYPE.
 FT DNA BIND 33 52 H-T-H MOTIF (POTENTIAL).
 FT SEQUENCE 238 AA; 26855 MW; C464FADCFCE68DAA CRC64;
 Query Match 17.1%; Score 7; DB 1; Length 238;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 28 AREELSE 34
 Db 32 AREELSE 38
 RESULT 7
 FADR_YERPE STANDARD; PRT; 238 AA.
 ID FADR_YERPE
 AC Q8ZEL9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Fatty acid metabolism regulator protein.
 GN FADR OR YPO2144 OR Y2177.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,

Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
 Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
 Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
 "Genome sequence of *Yersinia pestis*, the causative agent of plague.";
 Nature 413:523-527(2001).
 RL Nature 413:523-527(2001).
 CC -1- FUNCTION: Multifunctional regulator of fatty acid metabolism.
 CC Represses transcription of at least eight genes required for fatty
 CC acid transport and beta-oxidation among which fadA, fadB, fadD,
 CC fadL and fadE. Activates transcription of at least three genes
 CC required for unsaturated fatty acid biosynthesis: fabA, fabB and
 CC iclR, the gene encoding the transcriptional regulator of the
 CC aceBAK operon encoding the glyoxylate shunt enzymes. Binding of
 CC fadR is specifically inhibited by long chain fatty acyl-CoA
 CC compounds (By similarity).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: Contains 1 HTH gntR-type DNA-binding domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ141451; CAC90954.1; -
 CC EMBL: AE013821; AAM85739.1; -
 CC PIR: AF0261; AF0261.
 CC HAMAP: MF_00696; -; 1.
 CC InterPro: IPR008920; FadR_C.
 CC InterPro: IPR000524; HTH_GntR.
 CC Pfam: PF00392; gntR; 1.
 CC PRINTS: PR00035; HTHGNTR.
 CC SMART: SM00345; HTH_GNTR; 1.
 CC PROSITE: PS00043; HTH_GNTR_FAMILY; 1.
 CC Fatty acid metabolism; Transcription regulation; Activator; Repressor;
 KW DNA-binding; Complete proteome.
 FT INIT MET 0 0 BY SIMILARITY.
 FT DOMAIN 1 70 HTH_GNTR-TYPE.
 FT DNA BIND 33 52 H-T-H MOTIF (POTENTIAL).
 FT SEQUENCE 238 AA; 27670 MW; 27C0885E976BCD3 CRC64;
 Query Match 17.1%; Score 7; DB 1; Length 238;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 28 AREELSE 34
 Db 32 AREELSE 38
 RESULT 8
 FADR_SHEON STANDARD; PRT; 240 AA.
 ID FADR_SHEON
 AC Q8ED80;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Fatty acid metabolism regulator protein.

```

GN FADR OR S02885;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297586; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imprim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Uterback T.R., McDonald L.A., C.M.;
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- FUNCTION: Multifunctional regulator of fatty acid metabolism (By
CC similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Contains 1 HTH GNTX-type DNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF015727; AAN55901.1; -
CC DR TIGR; S02885; -
CC DR HAMAP; MF_00596; -; 1.
CC DR InterPro; IPR008920; Fadr_C.
CC DR InterPro; IPR000524; HTH_GntR.
CC DR Pfam; PF00392; gntR; 1.
CC DR PRINTS; PR00035; HTHGNTX.
CC DR SMART; SM00345; HTH GNTX; 1.
CC DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
CC Fatty acid metabolism; Transcription regulation; Activator; Repressor;
CC DNA-binding; Complete proteome.
CC FT DOMAIN 1 71 HTH GNTR-TYPE.
CC FT DNA BIND 34 53 H-T-H MOTIF (POTENTIAL).
CC FT SEQUENCE 240 AA; 27227 MW; 0EA97D59D7A15EBA CRC64;
CC -----
Query Match 17.1%; Score 7; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28 AEREISE 34
DB 33 AEREISE 39
|||||
|||||

RESULT 9
FADR_VIECH
ID FADR_VIECH STANDARD; PRT; 279 AA.
AC QSK0J6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fatty acid metabolism regulator protein.
GN FADR OR VC1900.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=22450551; PubMed=12562793;
RA Herz K., Vimont S., Pagan E., Berche P.;
RT "Roles of NhaA, NhaB, and NhaD Na+/H+ antiporters in survival of
RT Vibrio cholerae in a saline environment."
RL J. Bacteriol. 185:1236-1244(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emdolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -!- FUNCTION: Multifunctional regulator of fatty acid metabolism (By
CC similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Contains 1 HTH GNTX-type DNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF489522; AAC37924.1; -
CC DR EMBL; AE004265; AAF95048.1; -
CC DR TIGR; VC1900; -
CC DR HAMAP; MF_00596; -; 1.
CC DR InterPro; IPR008920; Fadr_C.
CC DR InterPro; IPR000524; HTH_GntR.
CC DR Pfam; PF00392; gntR; 1.
CC DR PRINTS; PR00035; HTHGNTX.
CC DR SMART; SM00345; HTH GNTX; 1.
CC DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
CC Fatty acid metabolism; Transcription regulation; Activator; Repressor;
CC DNA-binding; Complete proteome.
CC FT DOMAIN 1 71 HTH GNTR-TYPE.
CC FT DNA BIND 34 53 H-T-H MOTIF (POTENTIAL).
CC FT SEQUENCE 279 AA; 31997 MW; 47D710A53CC4F5B1 CRC64;
CC -----
Query Match 17.1%; Score 7; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28 AEREISE 34
DB 33 AEREISE 39
|||||
|||||

RESULT 10
FADR_VIEBA
ID FADR_VIEBA STANDARD; PRT; 279 AA.
AC Q87N05;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fatty acid metabolism regulator protein.
GN FADR OR VP2071.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=RMD 2210633 / Serotype O3:k6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shingawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RL distinct from that of V. cholerae."
RL Lancet 361:743-749 (2003).
CC -1- FUNCTION: Multifunctional regulator of fatty acid metabolism (By
CC similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Contains 1 HTH gntR-type DNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP005080; BAC60334.1; -
DR HAMAP: MF_00696; -; 1.
DR InterPro: IPR008920; Fadr_C.
DR Pfam: PF00392; gntR; 1.
DR PRINTS: PR00035; HTHGNTR.
DR PROSITE: PS00043; HTH_GNTR_FAMILY; 1.
DR Fatty acid metabolism; Transcription regulation; Activator; Repressor;
KW DNA-binding; Complete proteome.
FT DOMAIN 1 71 HTH GNTR-TYPE.
FT DNA_BIND 34 53 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 279 AA; 32069 MW; 8429B92B45149399 CRC64;

Query Match 17.1%; Score 7; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AEREISE 34
Db 33 AEREISE 39

RESULT 11
FADR_VIBVU STANDARD; PRT; 279 AA.
AC Q9DAG8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fatty acid metabolism regulator protein.
GN FADR_OR VJ12233.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Multifunctional regulator of fatty acid metabolism (By
CC similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Contains 1 HTH gntR-type DNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE016804; AA010613.1; -
DR HAMAP: MF_00696; -; 1.
DR InterPro: IPR008920; Fadr_C.
DR Pfam: PF00392; gntR; 1.
DR PRINTS: PR00035; HTHGNTR.
DR SMART: SM00345; HTH_GNTR; 1.
DR PROSITE: PS00043; HTH_GNTR_FAMILY; 1.
DR Fatty acid metabolism; Transcription regulation; Activator; Repressor;
KW DNA-binding; Complete proteome.
FT DOMAIN 1 71 HTH GNTR-TYPE.
FT DNA_BIND 34 53 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 279 AA; 32072 MW; 733F71C9F167477 CRC64;

Query Match 17.1%; Score 7; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AEREISE 34
Db 33 AEREISE 39

RESULT 12
KHSE SULSO STANDARD; PRT; 311 AA.
AC Q97W70;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homoserine kinase (EC 2.7.1.39) (HK).
GN THRB OR SSO2367.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic V., Allard G.,
RA Aways M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasus G., Fletcher C., Gordon P.M.K., Medina N., Peng X.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Theriault C., Tolstrup N.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Duguet M., Gaasterland T.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-homoserine = ADP + O-phospho-L-
CC homoserine.
CC -1- PATHWAY: Threonine biosynthesis from aspartate; fourth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the GMP kinase family. Homoserine kinase
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006898; AA042518.1; -
DR HAMAP: MF_00384; -; 1.

```

```

DR InterPro: IPR006204; GHMP kinase.
DR InterPro: IPR006203; GHMPkinase ATP.
DR InterPro: IPR000870; Homoser_Kin.
DR Pfam: PF00288; GHMP_kinases; 1.
DR PRINTS: PRO0958; HOMSERKINASE.
DR TIGRFAMs: TIGR00191; thrB; 1.
DR PROSITE: PS00627; GHMP_KINASES ATP; 1.
KW Threonine biosynthesis; Transferase; Kinase; ATP-binding;
KW Complete proteome. 98 ATP (POTENTIAL).
FT NP_BIND 311 AA; 33649 MW; 4E39A20E5F83D507 CRC64;
SQ SEQUENCE 311 AA; 33649 MW; 4E39A20E5F83D507 CRC64;

Query Match 17.1%; Score 7; DB 1; Length 311;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 ESKTKA 28
DB 181 ESKTKA 187

RESULT 13
DP3X_BACSU STANDARD; PRT; 563 AA.
ID P09122;
AC P09122;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA polymerase III subunit gamma/tau (EC 2.7.7.7).
GN DNAX OR DNAX OR BSU00190.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9108245; PubMed=2124672;
RA Alonso C., Shirahige K., Ogasawara N.;
RT "Molecular cloning, genetic characterization and DNA sequence
RT analysis of the recM region of Bacillus subtilis.";
RL Nucleic Acids Res. 18:6771-6777(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Betero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschini C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton A.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Ettian K.D., Errington J., Fabret C., Ferrazi A., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ozawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Pressecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

```

```

RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Taconci E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256(1997).
RN [4]
RP SEQUENCE OF 1-422 FROM N.A.
RC STRAIN=168;
RX MEDLINE=88203213; PubMed=2452406;
RA Struck J.C.R., Vogel D.W., Ulbrich N., Erdmann V.A.;
RT "A dnazX-like open reading frame downstream from the Bacillus
RT subtilis scRNA gene.";
RL Nucleic Acids Res. 16:2720-2720(1988).
RN [5]
RP SEQUENCE OF 1-5 FROM N.A.
RC STRAIN=168;
RX MEDLINE=91002681; PubMed=1698458;
RA Struck J.C.R., Kretschmer-Kazemi F.R., Schroeder W., Hucho F.,
RA Toschka H.Y., Erdmann V.A.;
RT "Characterization of a 17 kDa protein gene upstream from the small
RT cytoplasmic RNA gene of Bacillus subtilis.";
RL Biochim. Biophys. Acta 1050:80-83(1990).
CC -!- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N)
CC -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC epsilon and theta chains) that associates with a tau subunit. This
CC core dimerizes to form the POLIII' complex. POLIII' associates
CC with the gamma complex (composed of gamma, delta, delta', psi and
CC chi chains) and with the beta chain to form the complete DNA
CC polymerase III complex (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X17014; CAA34877.1; -
DR EMBL: D26185; BAA05255.1; -
DR EMBL: Z99104; CAB11795.1; -
DR EMBL: X06803; CAA29958.1; -
DR EMBL: X52144; CAA36390.1; -
DR FIR: I40469; I40469.
DR FIR: S13786; S13786.
DR Subtilist; BG10083; dnax.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003959; AAA_ATPase_cent.
DR InterPro: IPR001270; Chaprin_c1pA/B.
DR InterPro: IPR008921; Polclamp_load_C.
DR InterPro: IPR000862; RFCdomain.
DR Pfam: PF00004; AAA; 1.
DR PRINTS: PRO0300; CIPPROTEASEA.
DR SMART: SM00382; AAA; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW ATP-binding; Complete proteome.
FT NP_BIND 45 211 H -> D (IN REF. 2 AND 3).
SQ CONFLICT 211 211
SQ SEQUENCE 563 AA; 62727 MW; 2BA428D32B8DC35C CRC64;

Query Match 17.1%; Score 7; DB 1; Length 563;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 RELSEQI 36

```


entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DB 315 RELSEQI 321

|||||

RESULT 14

MOES_HUMAN

ID MOES_HUMAN STANDARD; PRT; 576 AA.

AC P26038;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Moesin (Membrane-organizing extension spike protein).

GN MSN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1] SEQUENCE FROM N.A., AND SEQUENCE OF 1-15; 53-59 AND 413-434.

RP TISSUE=Placenta;

RC MEDLINE=92020840; PubMed=1924289;

RX Lankes W.T., Furthmayr H.;

RA "Moesin: a member of the protein 4.1-talin-ezrin family of proteins.";

RT Proc. Natl. Acad. Sci. U.S.A. 88:8297-8301(1991).

RN [2]

RP SEQUENCE FROM N.A.

RA Howden P.;

RA Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas;

RX MEDLINE=22398257; PubMed=12477932;

RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [4]

RP SEQUENCE OF 1-7.

RC TISSUE=Platelet;

RX MEDLINE=22608298; PubMed=12665801;

RA Gevaert K., Goethals M., Martens L., Van Damme J., Staes A.,

RA Thomas G.R., Vandekerckhove J.;

RT "Exploring proteomes and analyzing protein processing by mass

RT spectrometric identification of sorted N-terminal peptides.";

RL Nat. Biotechnol. 21:566-569(2003).

CC -!- FUNCTION: Probably involved in connections of major cytoskeletal

CC structures to the plasma membrane.

CC -!- TISSUE SPECIFICITY: In all tissues and cultured cells studied.

CC -!- SIMILARITY: Contains 1 FERM domain.

CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;

CC WWW=<http://www.infobiogen.fr/services/chronocancer/Genes/MSNID363.html>.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; M69066; AAA36322.1; -.

DR EMBL; Z98946; CAB46379.1; -.

DR EMBL; BC017293; AAH17293.1; -.

DR EMBL; A41289; A41289.

DR PDB; 1E5W; 27-JUN-01.

DR PDB; 1EF1; 10-MAY-00.

DR Aarhus/Ghent-2DPAGE; 3515; IEF.

DR Aarhus/Ghent-2DPAGE; 3516; IEF.

DR Genew; HGNC:7373; MSN.

DR MIM; 309845; -.

DR GO; GO:0005856; C:cytoskeleton; TAS.

DR GO; GO:0005886; C:plasma membrane; TAS.

DR GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.

DR InterPro; IPR000299; Band 4.1.

DR InterPro; IPR000798; Ez/rad/moesin.

DR InterPro; IPR008954; Moesin.

DR Pfam; PF00373; Band 4.1.

DR Pfam; PF00769; ERM_1.

DR PRINTS; PR00935; BAND4.1.

DR SMART; SM00295; B4.1; 1.

DR PROSITE; PS00660; FERM_1; 1.

DR PROSITE; PS00661; FERM_2; 1.

DR PROSITE; PS00657; FERM_3; 1.

XX Structural protein; Cytoskeleton; 3D-structure.

FT INIT MET 0 0

FT DOMAIN 1 294 FERM.

SQ SEQUENCE 576 AA; 67689 MW; 96CE7C28D55A8176 CRC64;

Query Match 17.1%; Score 7; DB 1; Length 576;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 REKEELM 15

Db 341 REKEELM 347

RESULT 15

MOES_MOUSE

ID MOES_MOUSE STANDARD; PRT; 576 AA.

AC P26041;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Moesin (Membrane-organizing extension spike protein).

GN MSN.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93055012; PubMed=1429901;

RA Sato N., Funayama N., Nagafuchi A., Yonemura S., Tsukita S.,

RA Tsukita S.;

RT "A gene family consisting of ezrin, radixin and moesin. Its specific

RT localization at actin filament/plasma membrane association sites.";

RL J. Cell Sci. 103:131-143(1992).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=NOD; TISSUE=Thymus;

RX MEDLINE=22354683; PubMed=12466851;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

RA Niki K., Osato N., Saito R., Suzuki H., Yamahata I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schreier L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shinada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Tesdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 11-576 FROM N.A.
RX MEDLINE=92243764; PubMed=1573844;
RA Furthmayr H., Lankes W.T., Amieva M.R.;
RT "Moesin, a new cytoskeletal protein and constituent of filopodia: its
RL role in cellular functions";
RL Kidney Int. 41:665-670(1992).
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal
CC structures to the plasma membrane.
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S47577; AAA1162.1; -
DR EMBL; AK088336; BAC40290.1; -
DR EMBL; BC047366; AAH47366.1; -
DR EMBL; M86390; AAA39728.1; -
DR MGD; MGI:91167; Msn.
DR InterPro; IPR000299; Band_4.1.
DR InterPro; IPR000798; Ez_rad/moesin.
DR InterPro; IPR008954; Moesin.
DR Pfam; PF00373; Band_41; 1.

DR Pfam; PF00769; ERM; 1.
DR PRINTS; PR00935; BAND41.
DR SMART; SMO0295; B41; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS00557; FERM_3; 1.
KW Structural protein; Cytoskeleton.
FT INIT MET 0 BY SIMILARITY.
FT DOMAIN 1 294 FERM.
FT CONFLICT 330 331 EL -> DV (IN REF. 4).
FT CONFLICT 370 371 RA -> SP (IN REF. 4).
SQ SEQUENCE 576 AA; 67635 MW; 5E0F4555552E9145 CRC64;

Query Match 17.1%; Score 7; DB 1; Length 576;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 9 REKEELM 15
DB 341 REKEELM 347

Search completed: August 16, 2004, 09:55:16
Job time : 33 secs

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	20	48.8	158	4	Q9UJZ7	Q9ujz7 homo sapien
2	20	48.8	159	4	Q9UJZ6	Q9ujz6 homo sapien
3	19	46.3	156	4	Q9UJZ8	Q9ujz8 homo sapien
4	13	46.3	161	4	Q9UKZ0	Q9ukz0 homo sapien
5	18	43.9	152	4	Q9UJZ2	Q9ujz2 homo sapien
6	12	29.3	421	11	Q8CBU4	Q8cbu4 mus musculus
7	12	29.3	455	11	Q8VHK3	Q8vhk3 rattus norvegicus
8	19.5	54.7	16	Q89K21	Q89k21 bradyrhizobium	
9	8	19.5	578	6	Q9GL03	Q9gl03 sus scrofa
10	8	19.5	901	5	Q8SUW1	Q8suw1 encephalito
11	7	17.1	55	2	Q9ZEW2	Q9zew2 enterobacte
12	7	17.1	128	2	Q5S576	Q5s576 vibrio algi
13	7	17.1	142	10	Q9C673	Q9c673 arabidopsi
14	7	17.1	146	4	Q9BRX7	Q9brx7 homo sapien
15	7	17.1	181	17	Q975A1	Q975a1 sulfolobus
16	7	17.1	195	4	Q8WV17	Q8wv17 homo sapien

RESULT 2
O9UJZ6

ID Q9UJ26 PRELIMINARY; PRT; 159 AA.
AC Q9UJ26;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Ezrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fadiel A., Chen Z.C., Naftolin F.;
RT "Mutation of ezrin gene in cancer.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF189213; AAF03156.1; -.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band 4.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR Pfam; PF00769; ERM_1.
DR PROSITE; PSS0057; FERM_3; 1.
FT NON_TER 1
FT NON_TER 159
SQ SEQUENCE 159 AA; 19234 MW; 7C398388B7BA70FA CRC64;

Query Match 48.8%; Score 20; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 2.6e-12; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 EREKEQMMREKEELMLRLQ 20
|||||
Db 137 EREKEQMMREKEELMLRLQ 156

RESULT 3
Q9UJ28 PRELIMINARY; PRT; 156 AA.
ID Q9UJ28;
AC Q9UJ28;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Ezrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Z.C., Fadiel A., Naftolin F.;
RT "Mutation of ezrin gene in cancer.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF188896; AAF03154.1; -.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band 4.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR Pfam; PF00769; ERM_1.
DR PROSITE; PSS0057; FERM_3; 1.
FT NON_TER 1
FT NON_TER 156
SQ SEQUENCE 156 AA; 19042 MW; 159F5AA684A2C3A6 CRC64;

Query Match 46.3%; Score 19; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 2.6e-11; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

Qy 1 EREKEQMMREKEELMLRLQ 19
|||||
Db 138 EREKEQMMREKEELMLRLQ 156

RESULT 4
Q9UK20 PRELIMINARY; PRT; 161 AA.
AC Q9UK20;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Ezrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Z.C., Fadiel A., Naftolin F.;
RT "Mutation analysis of ezrin gene in cancer cells.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187552; AAD56713.1; -.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band 4.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR Pfam; PF00769; ERM_1.
DR PROSITE; PSS0057; FERM_3; 1.
FT NON_TER 1
FT NON_TER 161
SQ SEQUENCE 161 AA; 19439 MW; 5FD6EB910E017099 CRC64;

Query Match 46.3%; Score 19; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.7e-11; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

Qy 1 EREKEQMMREKEELMLRLQ 19
|||||
Db 143 EREKEQMMREKEELMLRLQ 161

RESULT 5
Q9UJ22 PRELIMINARY; PRT; 152 AA.
ID Q9UJ22
AC Q9UJ22;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Cyrovillin 2 (Fragment).
GN VIL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Z.C., Fadiel A., Naftolin F.;
RT "Mutation of cyrovillin gene in glioblastoma.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF190059; AAF03158.1; -.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band 4.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR Pfam; PF00769; ERM_1.
DR PROSITE; PSS0057; FERM_3; 1.
FT NON_TER 1
FT NON_TER 152
SQ SEQUENCE 152 AA; 18416 MW; B7FCE817DD2B9291 CRC64;

Query Match 43.9%; Score 18; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 2.6e-10; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 0;

Qy 1 EREKEQMMREKEELMLRL 18
|||||
Db 134 EREKEQMMREKEELMLRL 151

```

FT  NON TER      455      455
SQ  SEQUENCE     455 AA;  54174 MW;  1FC9A95F4C7D5893 CRC64;

Query Match      29.3%; Score 12; DB 11; Length 455;
Best Local Similarity 100.0%; Pred.No. 0.00067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  9 REKEELMLRLQD 20
    |||||
Db  342 REKEELMLRLQD 353

RESULT 6
Q8CBU4 ID Q8CBU4 PRELIMINARY; PRT; 421 AA.
AC Q8CBU4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Villin 2 (Fragment).
GN Vill2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK035271; BAC29009.1; -.
DR MGD; MGI:98931; Vil2.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band 4.1.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR InterPro; IPR008954; Moesin.
DR Pfam; PF00373; Band_41; 1.
DR Pfam; PF00769; ERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS50057; FERM_3; 1.
FT  NON TER      1
SQ  SEQUENCE     421 AA;  50273 MW;  E8D8CD0D3F6E7497 CRC64;

Query Match      29.3%; Score 12; DB 11; Length 421;
Best Local Similarity 100.0%; Pred.No. 0.00063;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  9 REKEELMLRLQD 20
    |||||
Db  177 REKEELMLRLQD 188

RESULT 7
Q8VHK3 ID Q8VHK3 PRELIMINARY; PRT; 455 AA.
AC Q8VHK3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ezrin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Gunn-Moore F.J.; Tait S.; Brophy P.J.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF450298; AAL47844.1; -.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band_4.1.1.
DR InterPro; IPR000798; Ez/rad/moesin.
DR Pfam; PF00373; Band_41; 1.
DR Pfam; PF00769; ERM; 1.
DR PRINTS; PR00935; BAND41.
DR SMART; SMC0295; B41; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS50057; FERM_3; 1.

FT  NON TER      455      455
SQ  SEQUENCE     455 AA;  54174 MW;  1FC9A95F4C7D5893 CRC64;

Query Match      29.3%; Score 12; DB 11; Length 455;
Best Local Similarity 100.0%; Pred.No. 0.00067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  9 REKEELMLRLQD 20
    |||||
Db  342 REKEELMLRLQD 353

RESULT 8
Q89K21 ID Q89K21 PRELIMINARY; PRT; 547 AA.
AC Q89K21;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Acyl CoA dehydrogenase.
GN AIDB OR BL15097.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005953; BAC50362.1; -.
DR GO; GO:0003995; F:acyl-CoA dehydrogenase activity; IEA.
DR GO; GO:0016431; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006089; Acyl-CoA dh.
DR InterPro; IPR006090; Acyl-CoA dh.C.
DR InterPro; IPR006091; Acyl-CoA dh.M.
DR Pfam; PF00441; Acyl-CoA dh; 1.
DR Pfam; PF02770; Acyl-CoA dh.M; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
KW Complete proteome.
SQ  SEQUENCE     547 AA;  59318 MW;  62C18C6DDA7B1F6F CRC64;

Query Match      19.5%; Score 8; DB 16; Length 547;
Best Local Similarity 100.0%; Pred.No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  27 KAERELSE 34
    |||||
Db  41 KAERELSE 48

RESULT 9
Q9GL03 ID Q9GL03 PRELIMINARY; PRT; 578 AA.
AC Q9GL03;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Period protein (Fragment).
GN PER1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
```

RP SEQUENCE FROM N.A.
RC STRAIN=F1 Bear Large White/Weishan;
RA Skinner T.M., Anderson S.I., Lopez-Corralles N.L., Loudon A.S.,
RA Haley C.S., Archibald A.L.;
RT "Genetic and physical mapping, expression analysis and partial gene
sequence of porcine Peri.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277735; CAC15393.1; -;
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR0001610; PAS.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00989; PAS; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS01112; PAS; 1.
FT NON TER 578 578
SQ SEQUENCE 578 AA; 62238 MW; 9A80CF55895E32AB CRC64;
Query Match 19.5%; Score 8; DB 6; Length 578;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 31 ELSEQIQ 38
DB 488 ELSEQIQ 495
RESULT 10
Q8SUW1 PRELIMINARY; PRT; 901 AA.
ID Q8SUW1
AC Q8SUW1
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein ECU07_1530.
GN ECU07_1530.
OS Encephalitozoon cuniculi.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RC MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prensier G., Barbe V., Peyretallade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
Encephalitozoon cuniculi.";
RL Nature 414:450-453 (2001).
DR EMBL; AL590447; CAD25684.1; -;
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW Ribosomal protein; Hypothetical protein.
SQ SEQUENCE 901 AA; 96503 MW; 37C3D28A2F3A92 CRC64;
Query Match 19.5%; Score 8; DB 5; Length 901;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22 EEXTKKA 29
DB 141 EEXTKKA 148
RESULT 11
Q9ZEW2

ID Q9ZEW2 PRELIMINARY; PRT; 65 AA.
AC Q9ZEW2
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE CopB (Fragment).
GN COPB.
OS Enterobacter intermedium.
OG Plasmid pLVI402.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Enterobacter.
OX NCBI_TaxID=61648;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=AH14;
RX MEDLINE=20431899; PubMed=10974114;
RA Osborn A.M., Da Silva Tatley F.M., Steyn L.M., Pickup R.W.,
RA Saunders J.R.;
RT "Mosaic plasmids and mosaic replicons: evolutionary lessons from the
analysis of genetic diversity in IncFII-related replicons.";
RL Microbiology 146:2267-2275 (2000).
DR EMBL; AJ009981; CAA08928.1; -;
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
FT NON TER 1 1
SQ SEQUENCE 65 AA; 7477 MW; 478B9CECFD6C6092 CRC64;
Query Match 17.1%; Score 7; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 ERELSEQ 35
DB 55 ERELSEQ 61
RESULT 12
Q56576 PRELIMINARY; PRT; 128 AA.
ID Q56576
AC Q56576
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Multifunctional regulator of fatty acid metabolism (Fragment).
GN FADRV.
OS Vibrio alginolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=663;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=B138-2;
RX MEDLINE=96332623; PubMed=8695633;
RA Nakamura I., Enomoto H., Unemoto T.;
RT "Cloning and sequencing of the mbaV gene, encoding an Na+/H+
antiporter from Vibrio alginolyticus.";
RL Biochim. Biophys. Acta 1275:157-160 (1996).
CC -1- SIMILARITY: BELONGS TO THE GNTF FAMILY OF TRANSCRIPTIONAL
REGULATORS.
DR EMBL; D83728; BAA12085.1; -;
DR GO; GO:0005822; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR008920; Fadr_C.
DR InterPro; IPR000524; HTH_GntR.
DR Pfam; PF00392; gntR; 1.
DR PRINTS; PR00035; HTHGNT.
DR SMART; SM00345; HTH_GNTF; 1.
DR PROSITE; PS00043; HTH_GNTF_FAMILY; 1.
KW DNA-binding; Transcription; Transcription regulation.
FT NON TER 128 128
SQ SEQUENCE 128 AA; 14443 MW; 2B529BCA40E15EDB CRC64;

```

Query Match      17.1%; Score 7; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 ABRELSE 34
DB 33 ABRELSE 39

RESULT 13
Q9C673 PRELIMINARY; PRT; 142 AA.
AC Q9C673;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN P28B23.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti T., Marziani A.,
RA Maltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL, AC079829; AGS0670.1; -.
DR PIR; C86389; C86389.
KW Hypothetical protein.
SQ SEQUENCE 142 AA; 17315 MW; 4DD3225FD6D609B7 CRC64;

Query Match      17.1%; Score 7; DB 10; Length 142;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 ETKKAE 29
DB 102 ETKKAE 108

RESULT 14
Q9BRX7 PRELIMINARY; PRT; 146 AA.
AC Q9BRX7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Histone deacetylase 6).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN SEQUENCE FROM N.A.
RP TISSUE=Brain, and Muscle;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor
RT vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005872; AAH05872.1; -.
DR EMBL; BC011498; AAH11498.1; -.
DR EMBL; BT006649; AAP35295.1; -.
DR InterPro; IPR000286; His_deacetylase.
DR Pfam; PF00850; Hist_deacetyl; 1.
KW Hypothetical protein.
SQ SEQUENCE 146 AA; 16399 MW; 6A2C1261CE1F83EB CRC64;

Query Match      17.1%; Score 7; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 EKEEML 16
DB 137 EKEEML 143

RESULT 15
Q975A1 PRELIMINARY; PRT; 181 AA.
AC Q975A1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein SF0512.
GN SF0512.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000982; BAB55506.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR SMART; SM00322; KH; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 181 AA; 20506 MW; 189F3BC6D019FF55 CRC64;

Query Match      17.1%; Score 7; DB 17; Length 181;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KAEREELS 33
DB 160 KAEREELS 166

Search completed: August 16, 2004, 09:56:03

```

Mon Aug 16 10:46:51 2004

us-09-856-070b-29.oli.rspt

Page 6

Job time : 37 secs